UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author PRESCOTT Katrina Rachel

An Analysis of the Molecular Basis of the 22q11 Deletion Syndrome

Dr Katrina Rachel Prescott

Molecular Medicine Unit
Institute of Child Health, London

A thesis submitted for the degree of Doctor of Philosophy to the University of London

2006

Supervisors: Professor Peter Scambler
Professor Robin Winter
Dr. Paul Riley

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Katrina Rachel Prescott confirm the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indication in the thesis.

27.06.06

Katrina Prescott

Abstract

The 22q11 deletion syndrome (22q11DS) is a developmental syndrome of the pharyngeal region. *Tbx1* haploinsufficiency is thought to account for the main structural anomalies observed in the 22q11 deletion syndrome. The "*Df1*" deletion in the mouse provides a model of 22q11DS, the deletion reflecting *Tbx1* haploinsufficiency in the context of the deletion of 21 adjacent genes.

The aims of this study were to examine genes within the *Tbx1* pathway. This was initiated by various experimental approaches including yeast two hybrid and Affymetrix expression microarray analysis.

Initially a yeast two hybrid screen was conducted using the mouse Tbx1 protein as bait, to determine potential protein binding partners for Tbx1. Using the N terminus of Tbx1 four independent clones of the cytoplasmic dynein light chain 1 protein were identified. This protein is a potential Tbx1 interactor and the implications of this and future follow-up work are discussed.

Secondly an Affymetrix microarray screen was conducted using RNA from pooled branchial arch dissections of E10.5 *Df1* and wild type embryos. *Df1* and wild type RNA were hybridised and the data were analysed using the Genespring statistical computer program. Potential upregulated and downregulated targets were identified and some were verified using real time quantitative PCR and whole mount in situ hybridisation. A similar study was performed using RNA from *Pax3* homozygous and heterozygous embryos to exam dysregulated genes from a mouse model with a similar phenotype to 22q11DS.

Finally to examine the possibility of related pathways implicated in 22q11DS pathogenesis an array-comparative genomic hybridisation (array-CGH) study was performed in collaborated with Prof Nigel Carter at the Sanger institute by analysing DNA from patients with features of 22q11DS, but no deletion within the DiGeorge region on chromosome 22q11. Three patients with chromosomal imbalances were identified and two were confirmed by FISH and microsatellite analysis.

Acknowledgements

I would like to thank my supervisors, Professor Peter Scambler and Dr Paul Riley for their guidance and excellent supervision of this project. I am particularly indebted to the late Professor Robin Winter who supervised me at the start of my PhD and I feel privileged to have had advice from such an enthusiastic tutor. I would like to thank my collaborators: Dr Deborah Henderson and Professor Andrew Copp for their help with the *Splotch* mice; Dr Kathryn Woodfine, Dr Charles Shaw-Smith and Prof Nigel Carter for the array-CGH work and Dr Mike Hubank for allowing me access to the Affymetrix array facililty. Supervision of the FISH work was kindly provided by Paula Stubbs and Dr Roger Palmer at the Great Ormond Steet hospital cytogenetics laboratory. Financial support was provided by a Wellcome Trust training fellowship.

I am grateful to so many members of the Molecular Medicine Unit who have helped me in different ways. I would like to particularly thank Dr Sarah Ivins who introduced me to laboratory work; Catherine Roberts, who guided me through the animal and embryo work; Dr Sugi Hussain for his help with the yeast two hybrid screen; Dr Chiara Bachelli who provided help with the microsatellite work and was a good friend throughout; and finally Dr Paris Ataliotis who was excellent in assisting me with everything from general crises through to computer work and provided important encouragement.

Table Of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Abbreviations

°C	degree Celsius
22q11DS	22q11 Deletion Syndrome
3-AT	3-aminotriazole
3D	3 dimensional
A	adenine
aa	amino acid
Ac	Acetate
ACTH	adrenal cortical releasing hormone
AD	Activation Domain
ADHD	attention deficit hyperactivity disorder
AHF	Anterior heart field
ANF	Atrial Natriuretic Factor
Apaf1	Apoptotic protease activation factor 1
APS	ammonium persulphate
Array-CGH	Microarray-comparative genomic hybridisation
ARSA	Anomalous right subclavian artery
ARVCF	Armadillo repeat gene deleted in VCFS
ASD	atrial septal defect
AVSD	Atrial ventricular septal defect
BAC	Bacterial artificial chromosome
Bad	Bcl-associated death promoter
BG	Background
β-gal	β-galactosidase
BH	Benjamini Hochberg
BIM	bcl-2 interacting protein
BLAST	basic local alignment search tool
BMP	bone morphogenic protein
Bok	Bcl-2 related ovarian killer protein
bp	base pair(s)
BSA	bovine serum albumin
cAMP	cyclic 3'5'-adenosine monophosphate
CASP3	Caspase 3, Apoptosis-Related Cysteine Protease
CDC45L	Cell Division Cycle 45, S. Cerevisiae, homologue-like
cDNA	complementary deoxyribonucleic acid
CES	cat eye syndrome
CFC	cryptic protein
CGH	comparative genomic hybridisation
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1propanesulfonate hydrate
CHD7	Chromodomain helicase DNA-binding protein 7
Cited2	CBP/p300-Interacting transactivator, with GLU/ASP-rich C-terminal domain, 2
CLTCT1	Clarithryn heavy polypeptide-like 1
CLTD	Clarithryn, heavy polypeptide D
cm	centimetre
cM	centiMorgan
CNS	central nervous system
COMT	Catechol-O-methyl transferase
CPX	Cleft Palate with Ankyloglossia
CREB	cAMP response element binding protein
CRK	V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog
CRKL	V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog-like
CRKOL	V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog-like
CT	Computerised tomography
Ct	Threshold cycle
CTAFS	conotruncal anomaly face syndrome
CTP	Citrate transport protein

CTP	Cytosine triphosphate
CUGBP2	CUG Triplet Repeat, RNA-Binding Protein 2
CX	Connexin
DAPI	4',6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
DBD	DNA binding domain
ddH_2O	distilled, deionised water
ddNTP	dideoxynucleoside triphosphate
Del	Deletion
DGCR	DiGeorge syndrome critical region
DGS	DiGeorge syndrome
DIG-dUTP	Digoxigenin labelled uridine triphosphate
DLC1	Dynein light chain 1, cytoplasmic
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNK	Data not known
dNTP	deoxynucleoside triphosphate
DORV	Double outlet right ventricle
dTTP	deoxythymidine triphosphate
Dup	Duplication
dUTP	deoxyuridine triphosphate
E(9.5)	mouse embryonic day (9.5)
E. coli	Escherichia coli
ECE1	Endothelin Converting Enzyme 1
ECG	Electrocardiogram
EDN	Endothelin
EDNRA	Endothelin receptor type A
EDNRB	Endothelin-B receptor
EDR1	Early development regulator 1
EDTA	ethylenediamine tetra-acetic acid
EEG	Electroencephalogram
ER	endoplasmic reticulum
ERK	extracellular signal related kinase
EST	expressed sequence tag
EtOH	Ethanol
EWG	Erect wing gene
FDR	False discovery rate
FGF	fibroblast growth factor
FISH	fluorescence *in situ* hybridisation
FOG	Friend of Gata
FOX	Forkhead box
g	gram
g	acceleration due to gravity
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA	Gata binding protein
GBX2	Gastrulation brain homeobox 2
GCS	Glucosylceramide Synthase
GNB1l	Guanine nucleotide binding protein
Gp1bβ	Glycoprotein 1b, platelet, beta polypeptide
GSCL	Goosecoid-like
GSK3β	glycogen synthase kinase 3β
GST	glutathione-S-transferase
HAND2	Heart and neural crest expressed derivatives 2
HDR	hypoparathyroidism, deafness, renal anomalies
Het	Heterozygous
Hic1	hypermethylated in cancer-1
HIRA	Histone cell cycle regulation defective, S. Cerevisiae, homologue of, A
Homo	Homozygous
HOS	Holt-Oram Syndrome
Hox	Homeobox

HTF9c	HpaIII tiny fragments locus 9c
IAA	Interrupted aortic arch
ICAT	b-catenin interacting protein 1
ICH	Institute of Child Health
IDD	Integral membrane protein deleted in DiGeorge syndrome
IQ	Intelligence Quotient
IU	international unit
IVT	In vitro transcription
JAG1	Jagged1
k	kilo
kb	kilobase pair
kDa	kiloDalton
l	litre
L	leucine
LCR	low copy repeat
LD	Learning Difficulties
LF	lunatic fringe
LiAc	Lithium Acetate
LiCl	Lithium Chloride
m	milli
M	molar; methionine
MA	Microarray
MAPCAs	multiple aorto pulmonary collateral arteries
MAPK	Mitogen activated protein kinase
Mb	megabase
Mdg1	microvascular differentiation gene 1
MEN1	Multiple endocrine neoplasia 1
MeOH	Methanol
MES	Methanesulfonic acid
Mins	Minutes
MITF	Microphthalmia-associated transcription factor
MLPA	Multiplex ligation dependent probe amplification
MM	Mismatch
MMP	Matrix metalloproteinase
MMU	Molecular Medicine Unit
mRNA	messenger RNA
MRPL40	Mitochondrial ribosomal Protein L40
MTC	Mutant testing correction
MW	molecular weight
n	nano
N-	amino
NADPH	Nicotinamide Adenosine Dinucleotide Phosphate Hydrogen
NCBI	National centre for biotechnology information
NCC	neural crest cell
NFKB1	Nuclear Factor Kappa-B, Subunit 1
NLVCF	Nuclear Localisation signal deleted in VCFS
No.	Number
NOS	Nitric Oxide synthase
NPDC1	Neural Proliferation and Differentiation Control gene 1
NRF-1	nuclear respiratory factor-1
NTD	Neural tube defect
OD	optical density
Oligo	oligonucleotide
OMD	Oxford Medical Database
OMIM	Online mendelian inheritance in man
p	pico
PA	pharyngeal arch
PAA	pharyngeal arch artery
PAGE	polyacrylamide gel electrophoresis
PA-VSD	pulmonary atresia with ventricular septal defect

Pax	Paired box gene
PBS	phosphate buffered saline
PCQAP	PC2 Glutamine/Q-rich associated protein
PCR	polymerase chain reaction
PDA	patent ductus arteriosus
PDGFRA	Platelet derived growth factor receptor alpha
PEG	polyethylene glycol
PIAS3	Protein Inhibitor of activated Stat3
PJS	Peter J. Scambler
PKA	Protein Kinase A
PKD	Polycystic kidney disease
PKD1	Polycystin-1
PM	Perfect Match
PNUTL1	Peanut like 1
POMC	pro-opiomelanocortin
PPARγ	peroxisome proliferator activated receptor γ
PPI	Prepulse Inhibition
PRODH	Proline dehydrogenase
PTA	persistent truncus arteriosus
PTH	Parathyroid hormone
PTPN11	Protein-tyrosine phosphatase, nonreceptor-Type, 11
RA	retinoic acid
RAA	right sided aortic arch
RALDH2	Retinaldehyde dehydrogenase 2
RAN	Ras related nuclear protein
RANBP1	RAN binding protein 1
RARE	retinoic acid response elements
RARG	Retinoic acid receptor γ
Refs	References
RMS	rhabdomysarcomas
RNA	Ribonucleic acid
Rpm	Revolutions per minute
RTN4R	Reticulin 4 receptor
RTQPCR	real time quantitative polymerase chain reaction
RXRα	Retinoid X receptor α
SAGE	Serial analysis of gene expression
SCARF	Scavenger receptor, class F, member 2
SCL25A1	Solute Carrier Family 25 member 1
SD	standard deviation from the mean
SD	Synthetic dextrose medium
SDCBP	Syndecan Binding Protein
SDS	sodium dodecyl sulphate
SEMA3A	Semaphorin 3a
SF	Scaling factor
SHH	Sonic hedgehog
SMM	supplemented minimal medium
SNAP-25	Synaptosomal associated protein-25
SNP	single nucleotide polymorphism
SOX	Sry-box
Sp	Splotch
SREC	Scavenger receptor expressed by endothelial cells
SSC	salt sodium citrate
STS	Sequence tagged Site
TAE	Tris-acetate/EDTA
TBE	Tris-borate/EDTA
TBX	T-Box
TDR	typically deleted region
TE	Tris/EDTA
TEMED	N,N,N'N'-tetramethylethylenediamine
TGA	Transposition of the great arteries

TMVCF	Transmembrane protein deleted in VCFS
TNF	Tumour necrosis factor
TOF	Tetralogy of Fallot
Tris	Tris(hydroxymethyl) amino-methane
tRNA	total RNA
TSK-2	Testes Specific Kinase
TUPLE1	Tup-like enhancer of split 1
TXNRD2	Thioredoxin reductase 2
U	unit; uracil
UAS	Upstream activation sequence
UFD1L	Ubiquitin Fusion Degradation like 1
UMS	Ulnar mammary syndrome
USP18	Ubiquitin specific protease 18
UTP	uridine triphosphate
UTR	Untranslated region
UV	ultraviolet
V	Volt
VCFS	Velo-Cardio-Facial Syndrome
VEGF	vascular endothelial growth factor
VPI	Velo-pharyngeal insufficiency
VSD	Ventricular septal defect
W	tryptophan
WS	Waardenburg syndrome
WT	Wild type
X-Gal	5-bromo-4-chloro-3-indolyl-D-galactoside
Y2H	Yeast Two hybrid
YAC	yeast artificial chromosome
ZDHHC8	Zinc finger and DHHC-domain containing protein 8
ZNF74	Zinc finger protein
μ	micro

An Analysis of the Molecular Basis of the 22q11 Deletion Syndrome

Chapter 1 Introduction

1.1 Historical Background

The 22q11 deletion syndrome (22q11DS) is characterised by conotruncal heart defects, hypocalcaemia and immunodeficiency (secondary to parathyroid and thymic hypoplasia respectively), palatal defects, learning difficulties and a distinct facial phenotype. Adults with this syndrome can develop major psychiatric illnesses including schizophrenia and bipolar affective disorder. It has an estimated prevalence of 1/4000 (Wilson et al. 1994) and in the majority of cases it is caused by a 3Mb (Megabase) deletion on chromosome 22q11.2 (Emanuel et al. 1998).

The syndrome has historically come under various different names. In 1955, Sedlačková, a speech therapist, published a series of 28 cases of a syndrome of congenitally short palate with hypernasal speech in the absence of overt palatal clefts. Sedlačková noted a distinctive facial appearance among the children including flaccid facial musculature (SEDLACKOVA 1955). The association between congenital thymic hypoplasia and hypocalcaemia was first described by Lobdell in 1959 (Lobdell 1959). In 1965 Angelo DiGeorge included the T cell dysfunction and cardiac defects (DiGeorge 1965). This led to the use of the name DiGeorge syndrome {Online Mendelian Inheritance in Man (OMIM) 601362} for patients with predominantly the immune defects. In 1972 Freedom *et al* noted the association with cardiac and facial anomalies (Freedom et al. 1972). Conotruncal anomaly face syndrome (CTAFS or Takao syndrome) was first reported by Kinouchi (Kinouchi et al. 1976) and links the conotruncal/facial features with developmental delay in a group of Japanese patients. In 1978 Robert Shprintzen and colleagues (Shprintzen et al. 1978) described the association between velopharyngeal incompetence/cleft palate, cardiac anomalies learning difficulties and a characteristic face and called it velo-cardio-facial syndrome (VCFS) or Shprintzen syndrome (OMIM 192430).

Originally all of these disorders were thought to be separate entities but it is now known that they encompass a spectrum of the same disorder termed 22q11DS in this review. The acronym CATCH 22 (*c*ardiac abnormalities, *a*bnormal facies, *t*hymic

hypoplasia, *c*left palate, *h*ypocalcaemia and deletion on chromosome *22*) has in the past been used to describe the syndrome, but is unacceptable due to its negative literary connotations (Joseph Heller's book *Catch-22* describes a no-win situation (Heller 1955)).

1.2 Clinical Characteristics of 22q11 Deletion Syndrome

The clinical phenotype of 22q11DS is extremely variable. As an example, there is a case report of monozygotic twins carrying the deletion with different cardiac anomalies (Goodship et al. 1995). This implies that stochastic and epigenetic events are also important in the development of the syndrome. This has been a limiting factor in early diagnosis and produces challenges to management and follow-up. Recognition of a patient with 22q11DS is important from an early age to allow for the diagnosis and treatment of the associated medical problems and to give accurate genetic counselling to the family. Table 1.2.1 shows the prevalence of some of the main clinical features seen in 22q11DS.

1.2.1 Craniofacial Features

Figure 1.2.1 shows a composite computer model of 46 patients with 22q11DS (Hammond et al. 2004). The subtlety of the facial features can be noted from the model. 90% of 22q11DS patients are reported to have the characteristic facial phenotype. 40% of patients have microcephaly and there are a few case reports of craniosynostosis (Dean et al. 1998).

The nose typically has a bulbous tip, sometimes with a midline groove and deficient alae nasae. The palpebral fissures are generally shorter than average. The jaw is retrognathic in 80% of cases. Many patients have malar hypoplasia and microstomia. 60-70% of patients have minor ear anomalies ranging from microtia to prominent "cup-shaped" ears with helical thickening.

Table 1.2.1 Prevalence of Clinical Features in 22q11DS

	Prevalence in 22q11DS	% of cases with 22q11 deletion
Cardiac Anomalies (overall)	**75%**[1]	
Ventricular septal defect (VSD)	15%[1]	10%[2]
Tetralogy of Fallot	20%[1]	10-15%[3]
Pulmonary atresia-VSD	10%[1]	20-50%[4]
Persistent truncus arteriosus	10%[1]	40%[5]
Interrupted aortic arch	15%[1]	20%[6]
Isolated aortic arch anomalies	5-10%[8]	24%[7]
Velopharyngeal		
Cleft Palate	14%[1]	2%[9]
Velopharyngeal insufficiency	32%[1]	Not studied/published
Pierre Robin Sequence	17%[10]	13%[10]
Ocular Abnormalities	**7%**[1]	**Not studied/published**
Central Nervous System		
Moderate-severe learning difficulties	18%[1]	Not studied/published
Hypotonia	70-80%	Not studied/published
ADHD	35-55%[11]	Not studied/published
Psychosis/Schizophrenia	20%[1]	2%[13]
Seizures	21%[1]	Not studied/published
Structural brain anomalies	3%[1]	Not studied/published
Cervical Spine anomalies	100%[12]	Not studied/published
Renal		
General Genital/Urinary anomalies	36%[1]	Not studied/published
Abnormal kidneys	17%[1]	Not studied/published
Obstructive anomalies	10%[1]	Not studied/published
Undescended testes	8%[1]	Not studied/published
Endocrine/Immunological		
Hypocalcaemia	60%[1]	Not studied/published
Growth Retardation (<3rd Centile height/weight)	36%[1]	Not studied/published
Growth hormone deficiency	4%[14]	Not studied/published
Aplastic thymus	14%[1]	Not studied/published

Reference Key

[1](Ryan et al. 1997);[2](Digilio et al. 1996);[3](McElhinney et al. 2003b);[4](Goodship et al. 1998);[5](McElhinney et al. 2003a);[6](Volpe et al. 2003);[7](McElhinney et al. 2001);[8](Momma et al. 1996);[9](Ruiter et al. 2003);[10](Goldberg et al. 1993);[11](Gothelf et al. 2004);[12](Ricchetti et al. 2004);[13](Lindsay et al. 1995);[14](Weinzimer & Driscoll 2001).

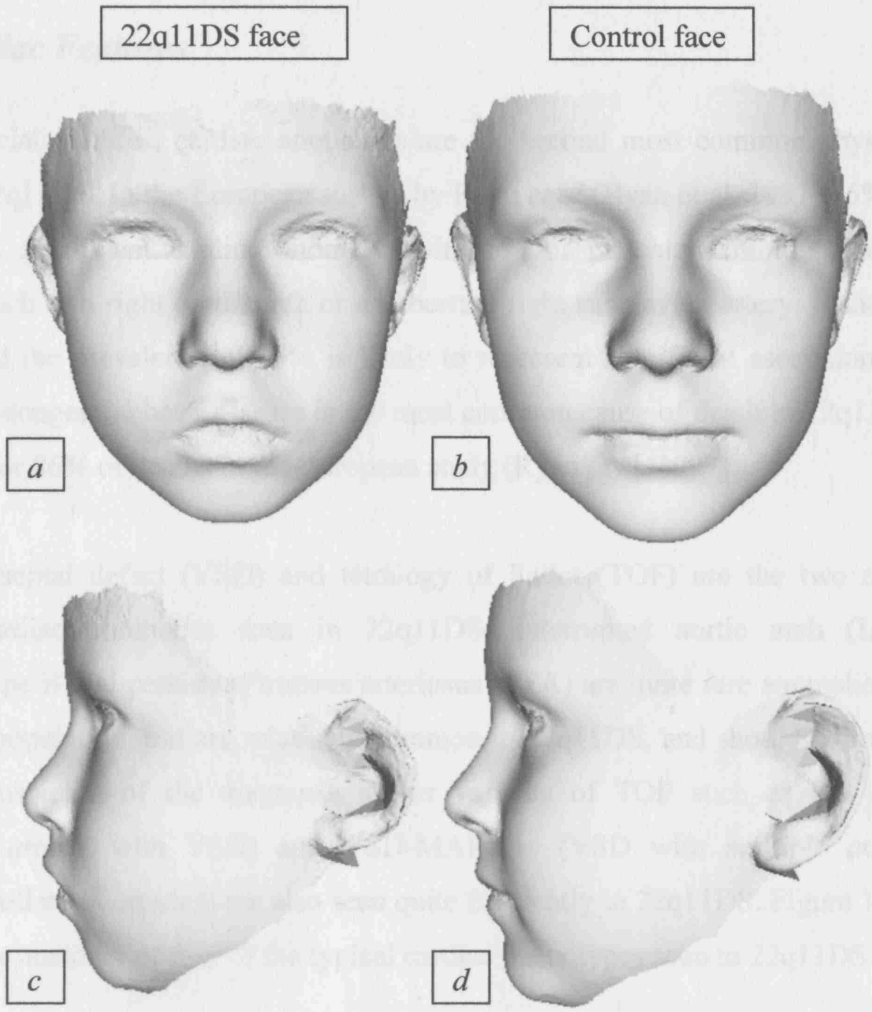


Figure 1.2.1 The 22q11DS facial phenotype.

The images above are derived from a set of 3D face scans of 46 individuals with 22q11DS and 129 controls between 2 weeks and 24 years. An overall mean face is generated using a dense correspondence of over 20,000 points. Then a dense surface model is used to interpolate faces between the average 22q11DS and control faces (Hammond et al. 2004). Figure 1.2.1*a* shows a 22q11DS portrait view obtained by interpolating as far as the 22q11DS face furthest from the mean. For comparison, the average control portrait is shown in 1.2.1*b*. Similarly computed 22q11DS and control profiles are shown in *c* and *d*. Note the overall face shape is narrower in *a* compared to *b*, and the palpebral fissures are shorter, the nose is narrower with a bulbous tip, the alae nasae and the mouth are smaller. By comparing *c* with *d* there is malar flattening with micrognathia and shallower orbits in the 22q11DS face (*c*). Images are courtesy of Prof. Peter Hammond, Eastman Dental Hospital, London.

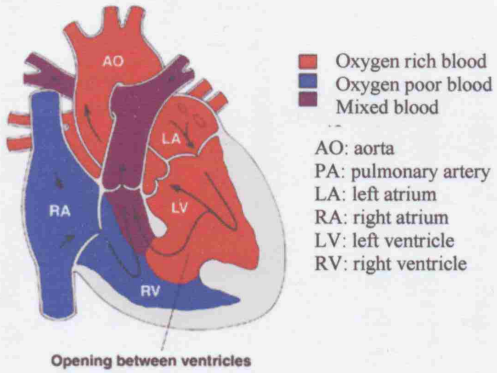
1.2.2 Cardiac Features

After the facial features, cardiac anomalies are the second most common physical finding in 22q11DS. In the European survey by Ryan *et al* (Ryan et al. 1997) 75% of cases had a significant cardiac anomaly, with 5% of patients exhibiting minor anomalies such as a right aortic arch or an aberrant right subclavian artery. It should be noted that the prevalence of 75% is likely to represent significant ascertainment bias. Severe congenital heart disease is the most common cause of death in 22q11DS accounting for 86% of deaths in the European study (Ryan et al. 1997).

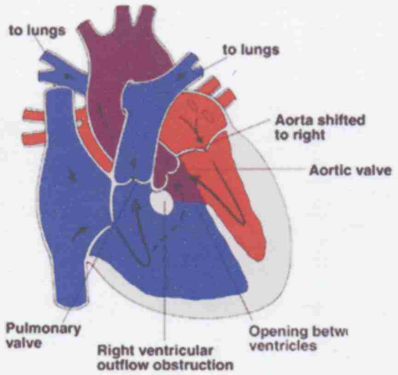
Ventricular septal defect (VSD) and tetralogy of Fallot (TOF) are the two most common cardiac anomalies seen in 22q11DS. Interrupted aortic arch (IAA) especially type B and persistent truncus arteriosus (PTA) are quite rare anomalies in the general population, but are relatively common in 22q11DS, and should therefore lead to a suspicion of the diagnosis. Other variants of TOF such as PA-VSD (*p*ulmonary *a*tresia with VSD) and VSD-MAPCAs (VSD with *m*ultiple *a*orto-*p*ulmonary *c*ollateral *a*rteries) are also seen quite frequently in 22q11DS. Figure 1.2.2 illustrates the anatomy of four of the typical cardiac phenotypes seen in 22q11DS.

A study in which 251 patients with conotruncal heart defects were examined for a 22q11 deletion, showed that the addition of vascular anomalies such as right sided aortic arch, significantly increased the chances of finding a 22q11 deletion (Goldmuntz et al. 1993). Patients with defects such as *d*ouble *o*utlet *r*ight *v*entricle (DORV) and *t*ransposition of the *g*reat *a*rteries (TGA) have only occasionally been reported in 22q11DS (Ryan et al. 1997).

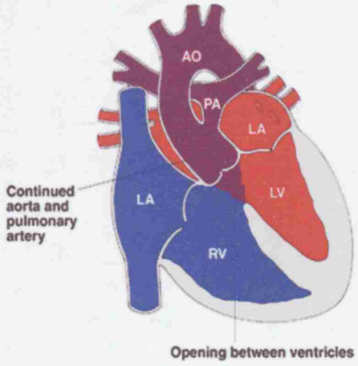
a) Ventricular Septal Defect



b) Tetralogy of Fallot



c) Persistent Truncus Arteriosus



d) Interrupted aortic arch type B

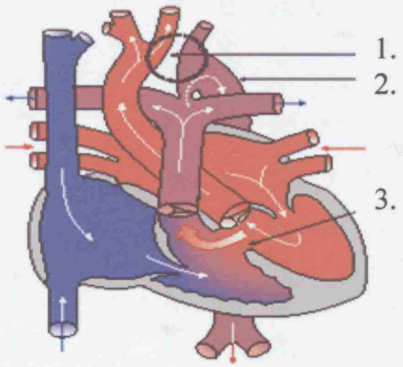


Figure 1.2.2 Typical cardiac anomalies seen in 22q11DS.

a) Illustration of the anatomy of a VSD showing mixing of deoxygenated and oxygenated blood through a common defect between the two ventricles.

b) Diagram of the anatomy of a Fallots defect where there is a VSD, an aorta that is overriding the two ventricles, right ventricular outflow tract obstruction and consequent right ventricular hypertrophy.

c) Depiction of a PTA where the pulmonary artery and the aorta are fused causing mixing of blood. There is commonly a VSD in surviving patients.

d) Illustration of an interrupted aortic arch type B
1. interruption of the aorta
2. the descending aorta receiving mixed blood from the PDA
3. a VSD where essential ventricular mixing occurs.

Figures adapted from http://www.schneiderchildrenshospital.org
.

1.2.3 Pharyngeal Anomalies

Defects affecting the pharyngeal region are highly significant as they impinge on breathing, feeding and communication. In the European study (Ryan et al. 1997) 14% of patients with 22q11DS were found to have a cleft palate and 32% velopharyngeal insufficiency (VPI). VPI is caused by poor motion of the lateral pharyngeal walls and/or the soft palate causing inadequate closure of the posterior pharynx during voice production and feeding. In infants this can lead to nasal regurgitation of feeds. In older children this can lead to a hypernasal voice and poor articulation, phonation and resonation. The diagnosis can be confirmed with videofluoroscopy studies and often pharyngoplasty is warranted to help speech and language development. In addition speech therapy combined with palatal surgery can often improve communication in affected children (Husein et al. 2004).

Cleft palates can be overt or submucous. Clefting of the lip is described in 2% of patients with 22q11DS and is usually unilateral if present (McDonald-McGinn et al. 1999). Feeding can often be difficult in babies with 22q11DS due to palatal anomalies and to lack of coordination of the oropharyngeal muscles.

Significant laryngeal and tracheal anomalies have been described in a proportion of patients with 22q11DS including laryngomalacia, laryngeal web, unilateral vocal fold paralysis, subglottic stenosis, tracheoesophageal fistula, tracheomalacia, laryngeal cleft and epiglottic anomalies (LeBlanc et al. 1996). In addition hearing loss due to high frequency sensorineural loss and chronic suppurative otitis media is relatively common in children with 22q11DS (Persson et al. 2003).

1.2.4 Developmental Profile

Recent extensive studies on the development of children with 22q11DS have shown a wide variation in intelligence with a mean Intelligence Quotient (IQ) of 73 +/-12.6 (Woodin et al. 2001). Severe mental retardation is rare and most children lie on the spectrum from moderate learning difficulties to average intelligence. Generally verbal IQ is noted to be higher than performance IQ (Woodin et al. 2001). Also reading skills tend to be considerably better than mathematical skills in 22q11DS

patients (Woodin et al. 2001). Verbal memory has been found to be an area of strength for such children, however visual-spatial memory and nonverbal reasoning is significantly below average (Bearden et al. 2001).

Many young patients have specific delay in areas of expressive and receptive speech and language. This is partly due to the physical anomalies described above and partly due to a cognitive/processing problem. In a recent study (Solot et al. 2001) it was found that 90% of 2 year olds with 22q11DS were nonverbal or just using single words by the age of 2 years. 80% of 3 year old were nonverbal or just using words or simple phrases. There was no correlation between language, speech, palate or cardiac findings in Solot's study (Solot et al. 2001).

1.2.5 Neurological Features

Generalised hypotonia can cause mild motor delay in 22q11DS. Coordination can also be a problem and children often need an evaluation by an occupational therapist and physiotherapist. Problems such as talipes equinovarus, scoliosis and leg pains are also more prevalent and can contribute to incoordination (Ryan et al. 1997).

Seizures occur in approximately 21% of patients with 22q11DS (Ryan et al. 1997). Most seizures are secondary to hypocalcaemia, emphasizing the need for early diagnosis and investigations in 22q11DS patients. Structural brain abnormalities are less common and occur in 3% of patients (Ryan et al. 1997). They include cerebellar vermal hypoplasia, cerebral atrophy, cerebral vascular anomalies, hydrocephalus, hypoplastic corpus callosum, cysts adjacent to the frontal horns, enlarged ventricles, small posterior fossa, polymicrogyria and neural tube defects. Analysis of the behavioural patterns in these patients has shown no relationship to these brain lesions (Mitnick et al. 1994).

1.2.6 Behavioural/Psychiatric Aspects

Behavioural problems are commonly found in patients with 22q11DS. Serious psychiatric illness in adult patients with 22q11DS is also highly prevalent. Studies

have shown that a significant number of children with 22q11DS have *a*ttention *d*eficit *h*yperactivity *d*isorder (ADHD) (Gothelf et al. 2004).

Psychiatric illnesses such as anxiety, depression, obsessions, compulsions and poor social interaction skills (Murphy 2002) are common in individuals with 22q11DS. Furthermore bipolar affective disorder has been reported in 36% of patients and psychosis in 10-29% (Paplos et al. 1996). In a large study by Murphy *et al* (Murphy et al. 1999), a psychotic disorder was seen in 30%, schizophrenia seen in 24% of patients and high rates of schizotypy were seen which can precede schizophrenia. The screening of 100 patients with schizophrenia yielded 2 patients with 22q11DS (Lindsay et al. 1995). In addition the prevalence of 22q11 deletions in childhood onset schizophrenia is higher than in the general population (Yan et al. 1998). Further prospective studies are underway to establish the relationship between 22q11DS and schizophrenia.

1.2.7 Genitourinary Anomalies

Renal anomalies were present in 36% of 22q11DS patients in the European study (Ryan et al. 1997). This was echoed by the study in Philedelphia (Wu et al. 2002) in which 31% were found to have anomalies such as renal agenesis or multicystic dysplastic kidney. 10% had obstructive anomalies of the urinary tract and 6% had vesicoureteric reflux. 6% of males had undescended testes and 8% hypospadius. Umbilical hernias occurred in 5% and inguinal hernias were reported in 20-30% of patients (Thomas & Graham, Jr. 1997).

1.2.8 Endocrine/Autoimmune Disease

Parathyroid hypoplasia can lead to hypocalcaemia in 22q11DS. This is more common in infants and can lead to seizures, tremors or tetany if untreated. Calcium homeostasis usually normalises if some parathyroid tissue is present and the glands hypertrophy. In the majority of cases the hypocalcaemia is transient and normalises as dietary calcium intake increases (Weinzimer & Driscoll 2001). More recent reports have identified symptomatic hypocalcaemia in older patients with the syndrome (Maalouf et al. 2004).

Other endocrine anomalies have been described in 22q11DS such as hypo and hyperthyroidism. Less than 1% of patients with 22q11DS develop hypothyroidism (Ryan et al. 1997). However autopsy reports have revealed multiple abnormalities of the isthmus and/or thyroid lobes in patients with 22q11DS (Freedom et al. 1972; Robinson 1975). Hyperthyroidism has been reported due to Graves disease in subjects with 22q11DS (Adachi et al. 1998), (Kawame et al. 2001). Autoimmune diseases affect 9% of patients with 22q11DS (Jawad et al. 2001) including juvenile rheumatoid arthritis (Rasmussen et al. 1996), Raynaud's phenomenon, autoimmune haemolytic anaemia, coeliac disease, diabetes and idiopathic thrombocytopenic purpura (Levy et al. 1997; Adachi et al. 1998), and are thought to result from the dysregulated immune system.

It is currently recommended that children with 22q11DS whose height or growth velocity is less than the 5th percentile for age should be tested for growth hormone deficiency (Weinzimer & Driscoll 2001). Children are more likely to show growth delay as a direct result of the syndrome, but in addition 4% of patients have growth hormone deficiency requiring treatment.

1.2.9 Immunological Deficits

As with most aspects of 22q11DS, the immune defects show wide clinical variability. Lymphopenia does not necessarily indicate severe infection susceptibility and vice versa. 80% of patients with 22q11DS have abnormalities of their immune system. As a consequence of thymic hypoplasia, a third of patients have decreased T cell numbers and are particularly vulnerable to infection in the first year of life. There is a tendency to experience recurrent sinopulmonary disease. In a small number of patients (<0.5%) there is complete thymic aplasia and absence of T cells (Ryan et al. 1997). The latter group are extremely vulnerable to severe infection and death and require prophylactic antibiotics, regular intravenous immunoglobulins and occasionally a thymic or bone marrow transplant.

In addition to the T cell defects B cell abnormalities and antibody deficiencies have also been described. IgA deficiency has been described in 2-10% of 22q11DS

patients and such patients have an increased incidence of autoimmune disorders (Gennery et al. 2002).

1.3 Clinical Genetic Implications

1.3.1 Selecting individuals for 22q11 deletion testing

As 22q11DS is such a heterogeneous disease it is recommended that a high index of suspicion is applied to patients presenting with some of the main features of 22q11DS. Although there are no published evidence based guidelines for 22q11DS testing, all patients with interrupted aortic arch and persistent truncus arteriosus should be considered for testing and cases of other cardiac anomalies should be considered for testing where there are additional features.

Table 1.3.1 provides a summary of the main syndromes with clinical overlap with 22q11DS. There is phenotypic overlap between 22q11DS, CHARGE and VACTERL syndrome as patients with 22q11 deletions have previously been described with coloboma and choanal atresia (both commonly found in CHARGE syndrome) and vertebral anomalies, radial defects and tracheoesophageal fistula (common features of VACTERL syndrome). Facially patients with Opitz G (BBB) look quite different to 22q11DS, but this syndrome has always been considered to have significant overlap because of the palatal anomalies, heart defects and feeding difficulties. Testing by fluorescence *in situ* hybridisation (FISH) for 10p15 deletions should be restricted to cases closely fitting the phenotype as although the facial features are very similar, deletions in this region are considerably rarer than 22q11 deletions (Peter Scambler-personal communication).

The majority of deletions arise *de novo*, but studies have shown that 10% of patients inherit the deletion from their parents (Ryan et al. 1997). Parents are often offered testing even if the deletion is not suspected clinically. A recent study where unselected relatives with deletion 22q11 were studied for their phenotype revealed only 40% of them had major findings which would have brought them to medical

attention (McDonald-McGinn et al. 2001). This illustrates further the range of phenotypic variability.

Testing of siblings is important in cases with parental deletions as 22q11DS is inherited in an autosomal dominant fashion. There are also incidences of germline mosaicism (Sandrin-Garcia et al. 2002) to consider when examining and testing siblings. The recurrence risk for a second child when a parent is affected is 50%, but for a second affected child in *de novo* cases a recurrence risk of 1% is generally quoted to allow for the possibility of germline mosaicism.

1.3.2 Prenatal Diagnosis of 22q11DS

Prenatal diagnosis of 22q11DS should be considered where an antenatal scan reveals a conotruncal heart defect. Sonographic demonstration of associated anomalies such as polyhydramnios, renal anomalies, cleft palate or skeletal anomalies may raise the suspicion of a 22q11 deletion in patients with cardiac defects. Preimplantation diagnosis using FISH has been performed in a patient with a 22q11 deletion (Iwarsson et al. 1998).

Table 1.3.1 Syndromes overlapping with 22q11DS

Syndromes with overlap (OMIM no)	**Features in common with 22q11DS**
CHARGE (214800)	Coloboma Heart Defects Atresia of Choanae Retardation of Growth/development Genital anomalies Ear anomalies (and deafness)
VACTERL Association (192350)	Vertebral Anomalies Anal Atresia Cardiac Anomalies TracheoEsophageal Fistula Radial defects Limb Anomalies
Partial Monosomy 10p (Distal:146255) (Proximal 601362)	**Distal deletion (involving GATA3)** HDR: Hypoparathyroidism Sensorineural Deafness Renal anomalies **Proximal Deletion** Cardiac anomalies Thymic hypo/aplasia
Opitz G (BBB) (145410)	Feeding difficulties in infancy Cleft palate Cardiac defects (DORV/VSD/RAA) Hypospadius/ Anal atresia Cerebral anomalies Malformed ears Learning difficulties
Cayler Syndrome (125520)	Asymmetrical crying facies Cardiac defects (VSD) Anal atresia
Vitamin A toxicity	Hypocalcaemia Microcephaly Cerebral anomalies Malformed ears Micrognathia Cleft palate Cardiac defects(PTA/VSD/Tetralogy of Fallot) Renal/genital abnormalities T-cell deficiency
Maternal Diabetes	Cardiac defects (VSD/PTA/ASD) Sacral agenesis/vertebral defects Renal agenesis
Fetal Alcohol Syndrome	Microcephaly/Learning difficulties/CNS abnormalities Ear anomalies Cardiac anomalies Cleft Palate (Tetralogy of Fallot/VSD/ASD) Renal anomalies

1.4 Aetiology of 22q11DS

1.4.1 Development of pharyngeal apparatus

The structures mainly affected in 22q11DS are derived from the pharyngeal/branchial arch region (figure 1.4.1). The pharyngeal apparatus is vertebrate specific with pharyngeal arches divided internally by endodermal pouches and externally by ectodermal clefts. Their development is complex involving coordination of cells from the ectoderm, endoderm, neural crest and mesoderm. Neural crest dependent and independent mechanisms are involved. All vertebrate embryos develop pharyngeal arches from a series of bulges on the lateral surface of the head. The ectoderm produces the epidermis and the sensory nerve ganglia. The pharyngeal endoderm of the pouches contributes to the development of the parathyroid, thymus and thyroid glands and gives rise to the epithelial cells lining the pharynx. The neural crest derived mesenchyme forms arch specific skeletal structures, the connective tissue, the smooth muscle cells of the arch arteries and the cells of the cardiac outflow tract. The paraxial mesoderm derived mesenchyme forms the muscular and endothelial precursors of each arch. Table 1.4.1 summarises the distinct structures formed by each arch.

The pharyngeal arch arteries (PAAs) run inside the pharyngeal arches, surrounded by mesenchymal cells. The aortic arch arteries develop in a craniocaudal direction forming a set of symmetrical arterial channels (figure 1.4.1b). The PAAs are initially paired and symmetric vessels that connect the outflow tract of the heart to the dorsal aortae. The first pair participates in forming the cranial vessels and largely disappears by the time the most caudal vessels have fully developed. The caudal PAAs (3^{rd}, 4^{th} and 6^{th}) remodel asymmetrically to form the mature aortic arch and its branches. The third arch arteries give rise to the left and right proximal common carotid arteries. The fourth aortic arches undergo extensive remodelling and form the preductal portion of the descending aorta on the left and the proximal right subclavian and brachiocephalic artery on the right. The fifth arch arteries regress and the sixth arch arteries undergo extensive remodelling and form the left ductus arteriosus and the proximal pulmonary arteries bilaterally. The remodelling process is important as it

has huge implication for the types of cardiac defects the patients have which in turn give clues to the developmental pathways involved.

Table 1.4.1 Table of the embryological origin of the pharyngeal structures

Germ Layer	1st Arch	2nd Arch	3rd Arch	4/5/6th Arch
Surface Ectoderm	External auditory meatus Tympanic membrane Buccal Cavity Jaw 1st pouch: middle ear	External auditory meatus Cervical sinus	Ear Epithelium	
Neural Ectoderm	V nerve	VII nerve	IX nerve	X/XI nerve
Aortic arches			Internal carotid artery	4th: aortic arch and right subclavian artery 6th: pulmonary arteries and PDA
Endoderm	Body of tongue Eustation tube	Thyroid Roof of tongue Tonsillar fossa	Roof of tongue Ventral thymus Parathyroid glands	Roof of tongue Epiglottis and pharynx
Mesoderm	Masticatory muscles Ear muscles	Muscles of facial expression	Stylopharyngeal muscle	Pharyngeal/laryngeal and palatal muscles
Skeleton	Maxilla Mandible Malleus/Incus	Stapes Styloid process Hyoid (lesser cornus)	Hyoid (greater cornus)	4th: thyroid cartilage Cricoid cartilage

Figure 1.4.1a. Diagram illustrating the contributions of neural crest cells to the outflow tract and the pharyngeal derivatives. Adapted from *Heart Development* by Margaret Kirby.

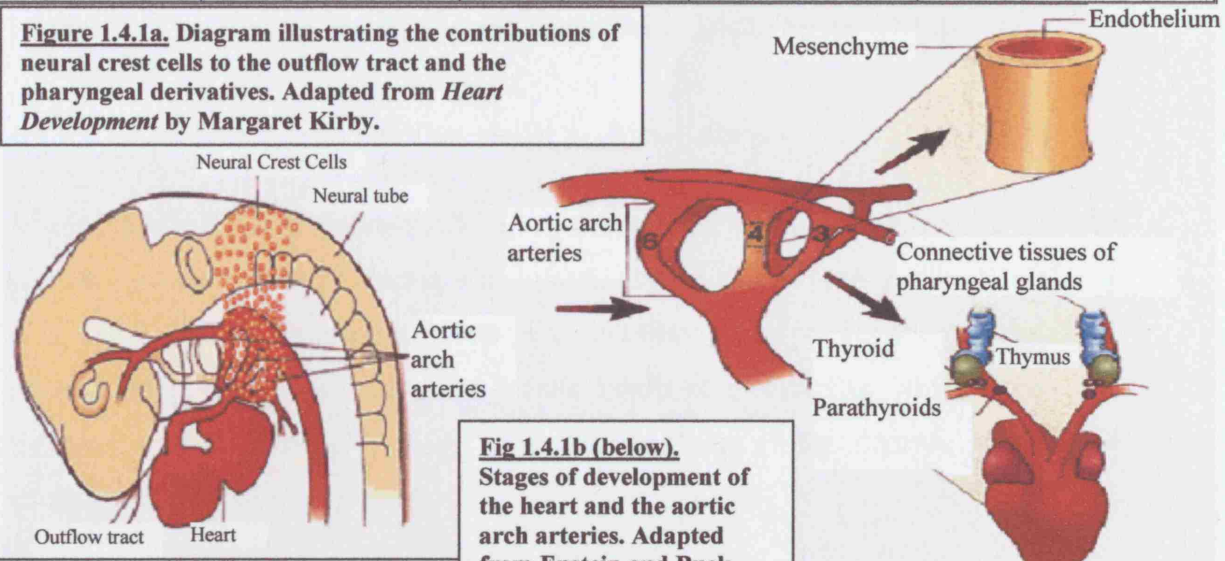
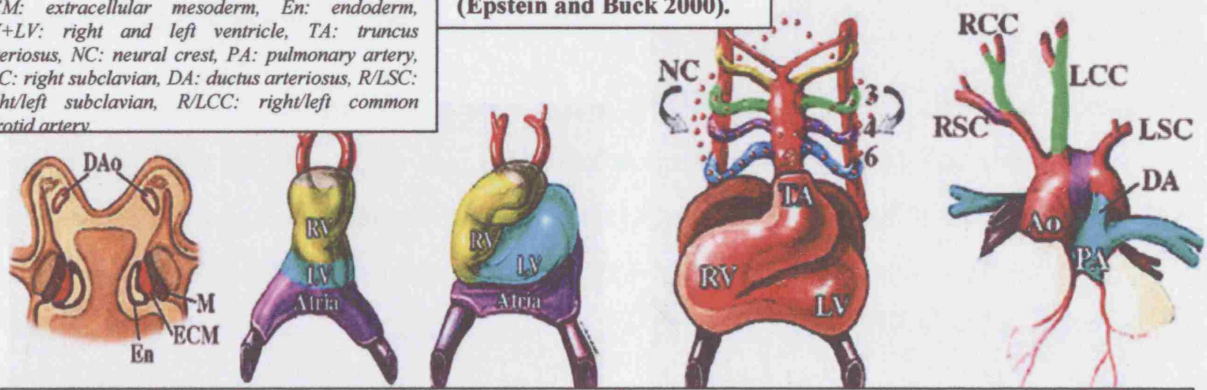


Fig 1.4.1b (below). Stages of development of the heart and the aortic arch arteries. Adapted from Epstein and Buck (Epstein and Buck 2000).

NC: neural crest. DAo: dorsal aortae, M: mesoderm, ECM: extracellular mesoderm, En: endoderm, RV+LV: right and left ventricle, TA: truncus arteriosus, NC: neural crest, PA: pulmonary artery, RSC: right subclavian, DA: ductus arteriosus, R/LSC: right/left subclavian, R/LCC: right/left common carotid artery



Human day 18 Mouse e7.5	day 22 e8.5	day 25 e10.5	day 34 e11.5	week 30 e18.5

1.4.1.1 The role of the neural crest

Numerous studies have emphasised the role of neural crest cells (NCCs) in patterning the pharyngeal arches. NCCs are essential for development of the aortic arch and outflow tract of the heart, craniofacial bones, the endocrine organs in the neck, the otic vesicle, the sympathetic and cranial nerve ganglia and for tooth development. NCCs migrate from the posterior and anterior hindbrain (rhombomeres 1 and 2) to populate the first pharyngeal arch and from the mid hindbrain region (rhombomeres 3 and 4) to fill the second arch. Post otic neural crest cells migrate from rhombomeres 6 and 7 to populate and pattern the third, fourth and sixth branchial arches. The smooth muscle cells of the arch arteries and the skeletal structures are also derived from neural crest cells. Some of the cells continue to migrate to the outflow tract where they become incorporated into the aortopulmonary septum and the conotruncal cushions where they contribute to septation of the truncus arteriosus. Neural crest cells also contribute to the formation of the thymus, the thyroid and parathyroid glands (figure 1.4.1a).

1.4.1.1.1 The neural crest ablation model in chick embryos

In 1983 Kirby *et al* performed ablation of the cardiac neural crest from the mid otic placode to somite three producing a phenotype similar to that seen in 22q11DS (Kirby et al. 1983). Heart defects were present in about 90% of embryos surviving to days 8-11 after ablation. The most prevalent defects involved the cardiac outflow tract and included DORV, TOF, PTA and VSD. Abnormalities of the thymus, thyroid and parathyroid glands were also seen.

1.4.1.2 The role of the endoderm

Despite evidence from the neural crest ablation model, it is also clear that the pharyngeal arches do form in the absence of neural crest (Graham 2001). The endoderm expresses signalling molecules such as bone morphogenetic proteins, fibroblast growth factors, sonic hedgehog, and transcription factors including *HoxA3, Pax1, Pax9 and Tbx1*. It is now thought that signals from the endoderm instruct NCCs and paraxial mesoderm to form mesenchyme.

Retinoid signalling is essential for patterning the pharyngeal endoderm. The formation of the second and more caudal pharyngeal arches is totally inhibited in vitamin-A

deficient quails (Chisaka & Capecchi 1991) which lack retinoids and in mice mutant for the *RALDH2* gene (Daw et al. 1996), which is necessary for the synthesis of retinoic acid.

1.4.2 The Genomics of 22q11DS

In 1981 de la Chapelle noted a 20:22 chromosomal translocation in a patient with features suggestive of 22q11DS (de la Chapelle et al. 1981). This prompted Scambler *et al* to establish that most patients with features of 22q11DS have a submicroscopic deletion on 22q11.2 making the region hemizygous (Scambler et al. 1992). Such deletions are now routinely tested for by FISH with the *HIRA/tuple* marker (OMIM 600237).

87% of patient have a 3Mb deletion on 22q11.2 and a further 7% have the same proximal breakpoint, but have a distal deletion endpoint extending only 1.5Mb (figure 1.4.2a) (Carlson et al. 1997; Ensenauer et al. 2003). The more common 3Mb deletion has been termed the *t*ypically *d*eleted *r*egion (TDR) and the nested 1.5Mb region is the *D*i*G*eorge *C*ritical *R*egion (DGCR). There are 27 genes within the DGCR (shown in figure 1.4.2.2) and 38 in the TDR (current map is best viewed via the ENSEMBL engine at http://www.ensembl.org/Homo_sapiens/mapview?chr=22). There is no evidence to suggest that the severity or the extent of the phenotype is affected by the size of the deletion (Matsuoka et al. 1998). There is no sex bias of the deleted chromosome indicating a lack of imprinting in the region (Morrow et al. 1995).

1.4.2.1 Low copy repeats on chromosome 22q11.2 produce recombination error hotspots

Deletion of the TDR on 22q11.2 is the most frequent interstitial chromosome deletion discovered in man based on current knowledge (Wilson et al. 1994). In addition to the two main types of deletion seen in 22q11DS there are three other main genomic disorders that map to the region (figure 1.4.2b). The first disorder is microduplication 22q11.2 (Dup 22q11). The phenotype shows some overlap with 22q11DS, although it is unclear if this is due to ascertainment bias as many patients are detected by chance when a FISH analysis for 22q11DS is requested. Conversely in a mouse model of 22q11DS, both increased and decreased gene dosage in the region orthologous to 22q11 has been shown to give a similar phenotype (Liao et al. 2004).

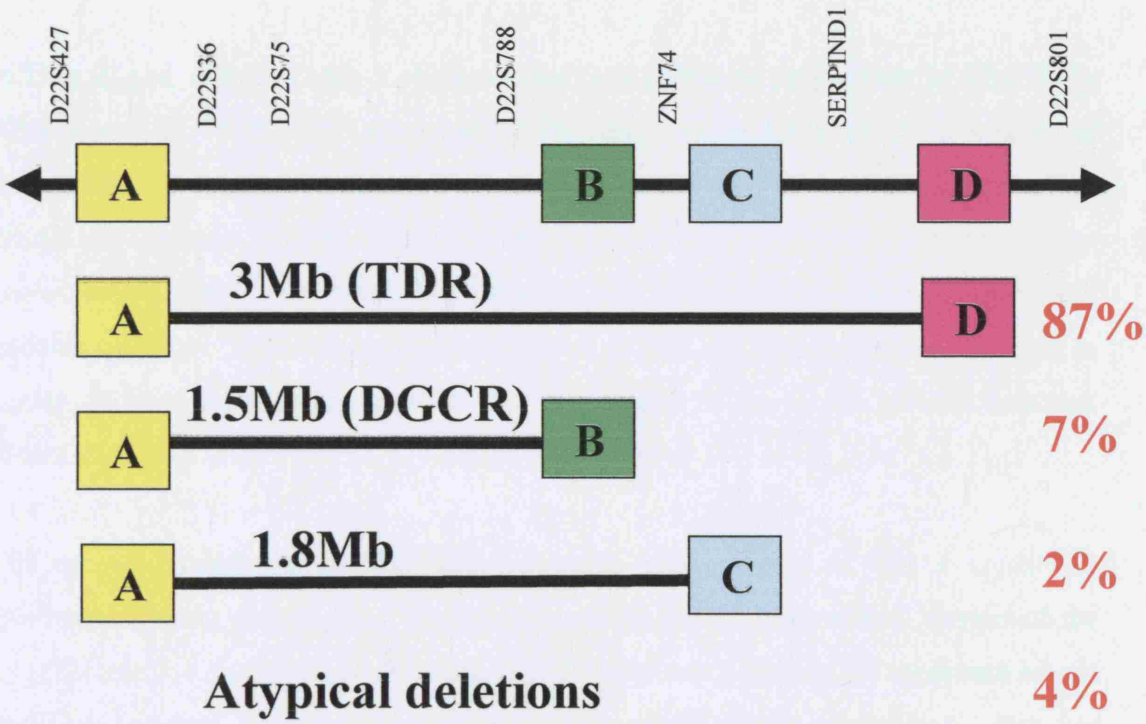


Figure 1.4.2a The 22q11 region including the locations of the LCR-22 segmental duplications A-D.

The names and relative locations of genes and STS markers are noted above the line. The figure is not drawn to scale. The percentages on the right give the relative percentages of patients that were found to be deleted by that given region by Saitta *et al*, 2004. **TDR**: typically deleted region. **DGCR**: DiGeorge syndrome Critical Region.

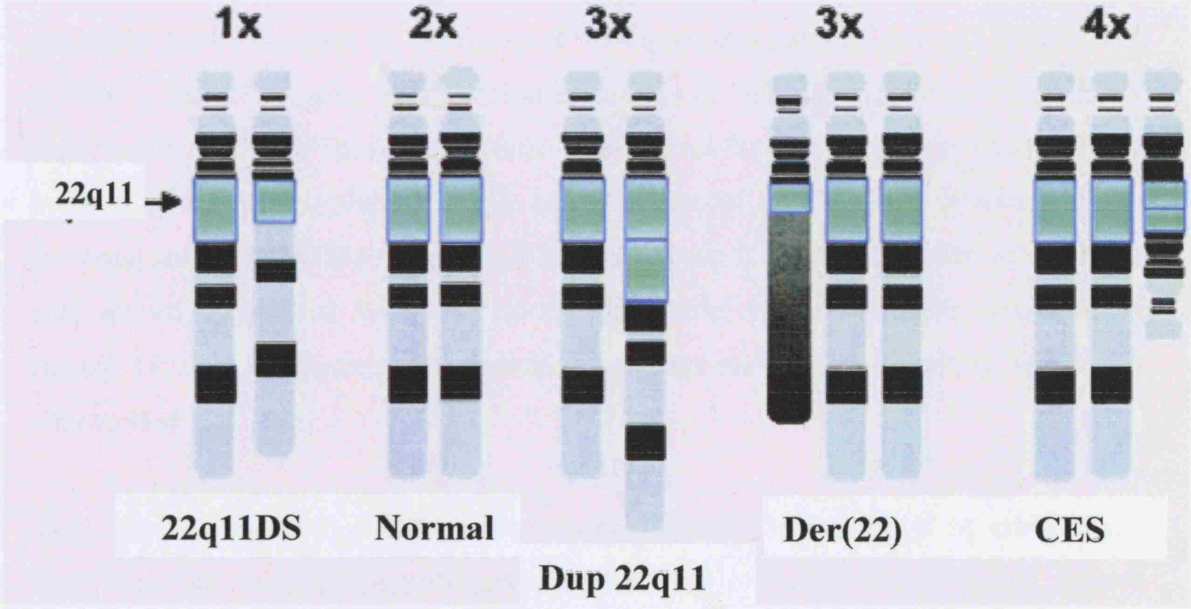


Figure 1.4.2b Chromosomal Copy number at 22q11.2

Illustrated above are the four main congenital anomaly disorders that are associated with germline rearrangements on 22q11 (boxed areas). The numbers above signify the dosage of the 22q11 region that is present in each. Adapted from McDermid and Morrow, 2002. **CES**: Cat Eye Syndrome.

In Dup 22q11, patients have a characteristic face (different to the face in 22q11DS), 70% have VPI, 38% cleft palate and 17% have congenital heart defects such as hypoplastic left heart, IAA-B and TOF (although numbers are so far relatively small). Occasional patients have been reported with absent thymus, intestinal malrotation (now known to be a relatively common feature of 22q11DS), urogenital abnormalities and variable cognitive impairment (Ensenauer et al. 2003). As duplications are difficult to detect cytogenetically, interphase FISH is the method of choice for optimal detection (Ensenauer et al. 2003).

The second disorder associated with increased gene dosage of 22q11 is der(22) syndrome. This syndrome occurs by nondisjunction events from normal carriers of the t(11;22) translocation (Zackai & Emanuel 1980). Patients with der(22) syndrome have a partial trisomy of both the 11q23-qter region and the 22q11-pter region. For this syndrome the t(11;22) breakpoint occurs in the same interval as the nested distal 1.5Mb 22q11DS breakpoint. Features of der(22) include heart defects (primarily atrial septal defects), renal malformations, cleft palate, preauricular tags and pits, anal atresia and learning difficulties (Knoll et al. 1995).

The third disorder is Cat Eye Syndrome (CES) (OMIM 115470) which is a rare disorder caused by a partial tetrasomy (4 copies) of the region that spans 22pter-q11 (Schinzel et al. 1981). The extra copies of this region are usually in the form of a chromosome that is supernumerary, bisatellited and dicentric called an invdup(22). There are two recurrent breakpoint regions involved in CES and these occur in the same intervals as the proximal and distal breakpoint of the TDR. Features of CES include ocular colobomata, anal atresia, congenital heart defects (typically total anomalous pulmonary venous return), renal malformations, pre-auricular skin tags and pits and learning difficulties (Schinzel et al. 1981).

Thus 22q11 is prone to a multitude of rearrangements (McDermid & Morrow 2002). Non allelic homologous recombination mediated by low copy repeat (LCR) regions result in the chromosomal rearrangement responsible for the syndromes (Edelmann et al. 1999; Shaikh et al. 2001). There are four main LCR regions on 22q11 (figure 1.4.2a). LCRs are genomic elements of approximately 200-500kb in length that share 94-99% homology with each other. During meiosis the LCRs can align inappropriately during recombination events, leading to a deletion on one chromosome and a duplication on

the other (figure 1.4.2.1). Both intra- and inter chromosomal homologous recombination events mediate the 3Mb deletion in 22q11DS. Inter-chromosomal recombination events occur between homologous chromosomes in meiosis I and are now thought to be the predominant mechanism of chromosomal exchange responsible for 22q11DS (Saitta et al. 2004).

Evidence for unequal meiotic exchanges in genomic regions containing LCRs has been provided for several human deletion syndromes such as Williams (7q11.2) (Bayes et al. 2003), Prader-Willi/Angelman (15q11.2) (Carrozzo et al. 1997), and Smith-Magenis syndrome (17p11.2) (Shaw et al. 2002). However the incidence of these microdeletion syndromes is considerably less than that for 22q11DS. It is postulated by Saitta *et al* that the region on 22q11 is predisposed to rearrangements as the LCRs are relatively large and densely packed onto the pericentric region, and that chromosome 22 is a small acrocentric chromosome (Saitta et al. 2004).

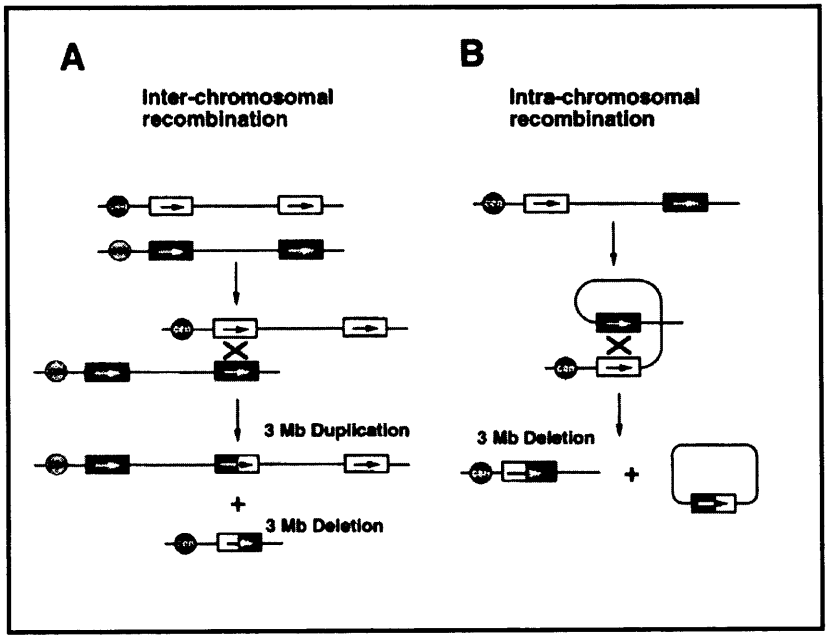


Figure 1.4.2.1 Mechanisms of crossover involving the LCRs

Mechanisms of production of the 3Mb deletions (A and B) and duplications (A) on chromosome 22q11.2. Diagram adapted from (Edelmann et al. 1999).

1.4.2.2 The search for a shortest region of overlap

Occasionally patients have been identified with atypical deletions involving the 22q11 region. In a study of 277 deleted patients with the 22q11DS phenotype by Saitta *et al* 6% of patients were identified without the typical 1.5Mb and 3Mb deletions (figure 1.4.2a) (Saitta et al. 2004). Figure 1.4.2.2 illustrates the location of many of the atypical deletions reported in the literature {the relevant reference follows the letter annotation in the figure 1.4.2.2; a-c:{Amati, 1999}, d:(Yamagishi et al. 1999), e:(Rauch et al. 1999), f:(O'Donnell et al. 1997), g:(Kurahashi et al. 1996), h:(McQuade et al. 1999), i:(Garcia-Minaur et al. 2002)}.

Gene mapping in some diseases has been assisted by the presence of a deleted region that is common to all patients (the shortest region of overlap). This can often narrow down a region within which the causative gene lies. In 22q11DS attempts to produce a shortest region of deletion overlap failed as some patients fell outside the DGCR (figure 1.4.2.2).

The study of patients with translocations has also been disappointing as no obvious candidate genes have been interrupted. This raised the possibility of the translocations and atypical deletions exerting long range positional effects on neighbouring genes (Schedl et al. 1996; Pfeifer et al. 1999). Phenotypic variability of 22q11DS could have led to some patients with atypical deletions having an overlapping but non-identical syndrome. It is also possible that a reduced dosage of different genes within the 22q11 region could produce similar syndromes.

1.4.3 Genes within the Typically Deleted Region

Several microdeletion syndromes are known where mutations within a single gene of non-deleted patients recapitulate the spectrum of defects observed in patients with a deletion. Examples include Rubenstein-Taybi syndrome (OMIM 180849) and Alagille syndrome (OMIM 118450). There are other syndromes such as Williams syndrome (OMIM 194050), where major features of a deletion syndrome can be ascribed to haploinsufficiency of specific contiguous genes within the deletion interval. It is still unclear in 22q11DS whether the entire spectrum of defects is the result of

haploinsufficiency of a single key gene within the TDR, or whether the full syndrome results from combined haploinsufficiency.

Candidate genes for the single gene theory have been chosen for study on the basis of their spacial and temporal expression pattern, their function and their mouse knockout phenotype. Table 1.4.3 shows the genes within the TDR region with a summary of what is currently known of their function. Genes marked with an asterix are within the DGCR.

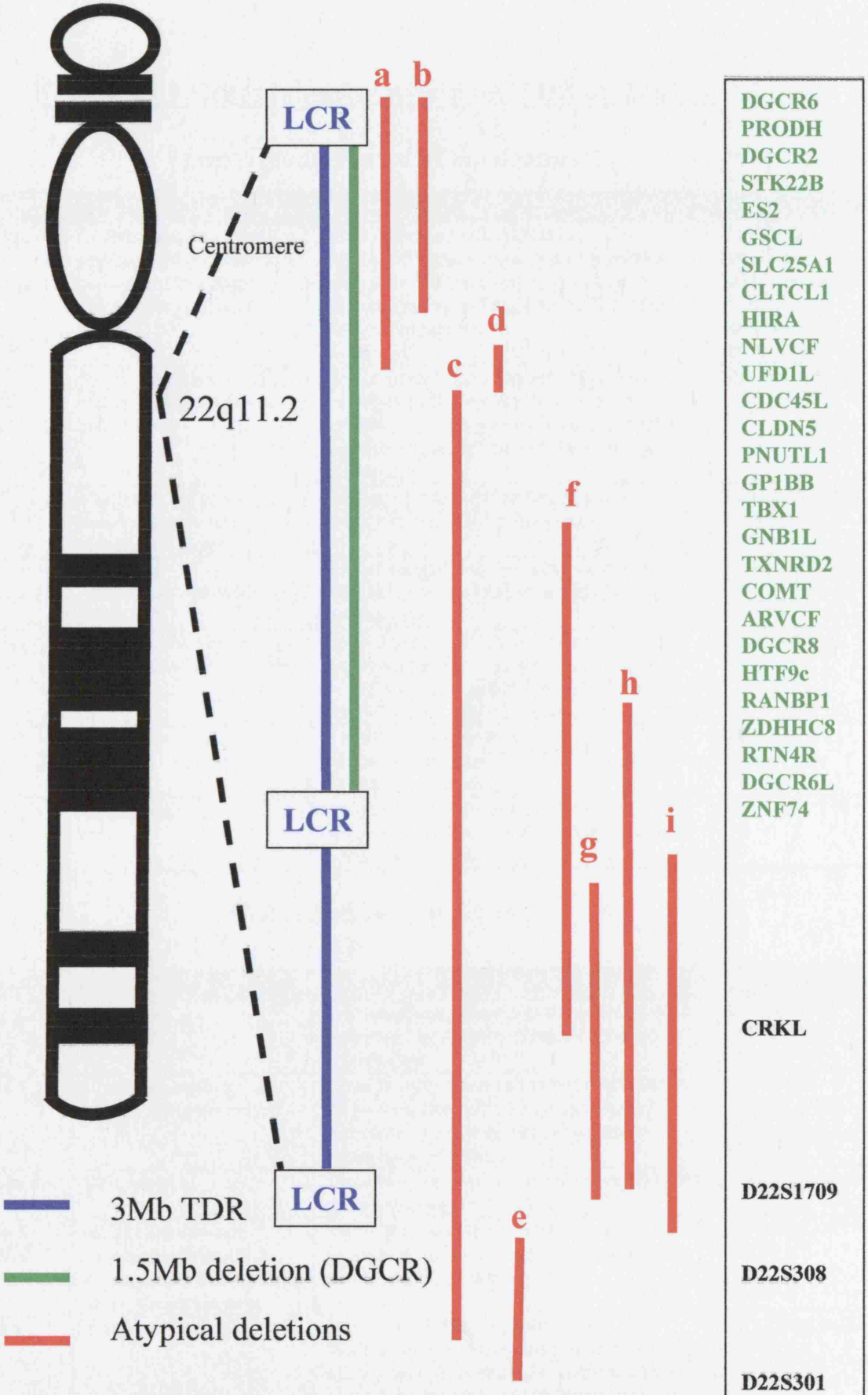


Figure 1.4.2.2 Composite breakpoint map of 22q11.

The common 3Mb and the nested 1.5Mb deletions are shown. Some reported atypical deletions are shown in red (see text for references of a-i). On the right are the genes within the TDR (in green). The location of 3 STS markers and CRKL are also shown. The low copy repeats (LCR) are shown in blue.

Table 1.4.3 Genes mapping within the TDR on 22q11.2

Transcription Factors/Coactivators

Gene	Expression	Function	References
TBX1*	Pharyngeal arches and pouches and otic vesicle, cardiac outflow tract. See text for details.	Main candidate for 22q11DS. Haploinsufficiency in the mouse produces aortic arch defects. Homozygous mice have other pharyngeal defects seen in 22q11DS. See text for details.	(Jerome & Papaioannou 2001; Lindsay et al. 2001; Merscher et al. 2001)
HIRA/ TUPLE1/ DGCR1*	Branchial arches 1 and 2, cranial neural folds, neural crest.	Transcription factor, interacts with *Pax3*. Possibly functioning as a cell cycle dependent repressor. Homozygote mice have abnormal heart looping and disturbed chamber formation.	(Roberts et al. 2002)
ZNF74*	Human neural crest derived tissues and foregut endoderm.	Zinc finger transcription finger with a possible role in RNA processing.	(Cote et al. 2001)
GSCL*	Branchial arches/pons/gut/testes.	Homeobox transcription factor related to goosecoid-required for craniofacial development	(Galili et al. 1998)
RANBP1/ HTF9A*	Forebrain, hindbrain, branchial arches, limb and heart.	Binding partner of RAN-involved in nucleocytoplasmic transport. Also required for the organization and function of the mitotic spindle in mammalian cells.	(Maynard et al. 2002)
ZDHHC8*		A putative transmembrane palmitoyltransferase. Candidate contributer to the risk of schizophrenia.	(Mukai et al. 2004)
PCQAP	Ubiquitous.	A subunit of a large protein complex which acts as a coactivator for RNA polymerase II driven transcription.	(Berti et al. 2001)

Cell Adhesion Molecules

Gene	Expression	Function	References
DGCR2/IDD/ LAN/Sez 12*	Ubiquitous.	Transmembrane protein-extracellular C-type lectin binding domain. Possibly involved in cell adhesion	(Budarf et al. 1995)
DGCR-6*	e7 mouse, brain/ spinal cord/pharyngeal arches.	Possibly involved in cell adhesion-similar to the gamma chain of laminin. Attenuated expression in chick embryos resulted in cardiovascular defects.	(Lindsay & Baldini 1997), (Hierck et al. 2004)
DGCR-6L*	Ubiquitous.	Thought to be involved in cell adhesion based on its homology to *DGCR6*.	(Edelmann et al. 2001)
Claudin 5/ TMVCF*	Cell-cell borders of endothelial cells. E9.5 mouse, adult human lung, heart, skeletal muscle.	Tight junction protein. Promotes activation of pro-matrix metalloproteinase-2 mediated by matrix metalloproteinases.	(Miyamori et al. 2001)
GP1Bβ*	Platelets.	Part of the platelet receptor for von Willebrand factor-mediating platelet adhesion and activation. Bernard Soulier syndrome can result from homozygous mutations and can therefore occur in 22q11DS. Hemizygosity can cause large platelets.	(Van Geet et al. 1998)
RTN4R*		Identified as a receptor for NOGO-66. Found to be mutated in patients with schizophrenia.	(McKerracher & Winton 2002; Sinibaldi et al. 2004)

Cell Cycle and Signalling

Gene	Expression	Function	References
CDC45L*		Homologue of yeast *CDC45* and involved in initiation of DNA replication.	(Yoshida et al. 2001)1
PNUTL1/ CDCREL1/ Septin 5*	Adult brain, heart and platelets.	Plays a role in cytokinesis. Co-purifies with SNAP-25-a possible role in synaptic vesicle transport, fusion or recycling events. Null mice appear normal.	(McKie et al. 1998)
UFD1L*	Ubiquitous.	Orthologue of yeast *UFD1*-required for ubiquitin dependent protein degradation. *Ufd1*+/- mice normal. Functional attenuation in chick embryos leads to cardiac OFT abnormalities.	(Rape et al. 2001), (Yamagishi et al. 2003a)
ARVCF*	Ubiquitous.	Functions at adherens junctions to control cell signalling through the *wnt* family of ligands and receptors. Morpholino knockdown in Xenopus results in disrupted gastrulation.	(Waibler et al. 2001), (Fang et al. 2004)
TSK-2 /STK22B*	Testes.	Serine/threonine kinase.	(Goldmuntz et al. 1997)
T10* D16H22S680 E in mouse	Ubiquitous.	Rich in serine/threonine residues. Activity possibly regulated by phosphorylation.	(Halford et al. 1993b)
CRKL Crkol in mouse		SH2-SH3-SH3 adaptor protein involved in growth and differentiation through cell signalling. Involved in the breakdown of adherens junctions. Homozygous mice have IAA defects and other neural crest derivatives. Heterozygote mice are normal.	(Feller 2001; Guris et al. 2001)
P2X6	CNS.	One of a family of purinoceptors mediating excitatory synaptic transmission and act presynaptically to modulate neurotransmitter release.	(Bobanovic et al. 2002)
SNAP29		A synaptosomal-associated protein likely to be involved in vesicle-membrane fusion processes such as neurotransmission and hormone secretion.	(Su et al. 2001)

Metabolic Cofactors and Enzymes

Gene	Expression	Function	References
TRXR2*	Mitochondrial.	Thioredoxin reductase-catalyzes the NADPH dependent reduction of the thioredoxin protein and protects against oxidative injury.	(Stoffel et al. 1996)
COMT*	Neuronal, placental, other tissues.	Catechol-O-methyl transferase breaks down neurotransmitters, catecholamines and drugs such as Levodopa. Implicated in schizophrenia.	(Grossman et al. 1992)
PRODH*		Mitochondrial enzyme-metabolises proline. Involved in the transfer of redox potential. Candidate for schizophrenia association in 22q11DS.	(Gogos et al. 1999)
SCL25A1/ CTP*	Ubiquitous.	Human mitochondrial citrate transporter protein.	(Heisterkamp et al. 1995)
PIK4CA		Phosphatidyinositol 4-kinase catalyses the first step in the biosynthesis of phosphatidylinositol 4,5 bisphosphate.	(Gehrmann & Heilmeyer, Jr. 1998)
SLC7A4/ CAT-4	Placental syncytiotrophoblast.	A cationic amino acid transporter.	(Ayuk et al. 2000)
HC-II		Heparin cofactor II precursor. Mutations are found in patients with pro-coagulant states.	(Baglin et al. 2002)

Genes with Miscellaneous Function

Gene	Expression	Function	References
HTF9C*		Shares the same promotor as Ranbp1. Similarity to some yeast and bacterial nucleic acid modifying enzymes. Function unknown.	(Bressan et al. 1991)
DGCR5		A possible pseudogene-similar to a seven pass transmembrane receptor.	(Sutherland et al. 1996)
DGCR8*		Contains 2 RNA binding domains and a WW domain. Required for MicroRNA processing.	(Gregory et al. 2004)
NLVCF/ MRPL40*	Ubiquitous-head, brachial arches 1 and 2.	Contains 2 consensus sequences for nuclear localisation signals.	(Funke et al. 1998)
GNB1L/ WDVCF*	Brain, forelimb, branchial arches, otic vesicle, somites, epithelium at the back of the eye.	Encodes a protein with a WD40 domain. Null mice are embryonic lethal.	(Gong et al. 2000 ; Ataliotis– personal communication)
DGSI /ES2/ DGCR14*	All tissues.	Homologous to yeast *Bis1*-interacts with proteins involved in RNA processing/modification.	(Taricani et al. 2002)
USP18/ UBP43		Homologous to ubiquitin-specific proteases. Involved in the innate immunity to viral infection.	(Malakhov et al. 2002), (Ritchie et al. 2004)
CLTCL1/ CLTD	Ubiquitous-high in skeletal muscle.	Clathrin heavy chain polypeptide. Possible role in receptor mediated endocytosis and signal transduction. Not found in the mouse genome.	(Gong et al. 1996)
SREC2/ SCARF2		Scavenger receptor expressed by endothelial cells. Interacts with *SREC1*.	(Ishii et al. 2002)

Over the last ten years various genes within the TDR have been in vogue for the ultimate accolade of the likely causative agent in 221q11DS. Some of these candidates will now be discussed.

1.4.3.1 HIRA

Historically, the first gene shown to be deleted in 22q11DS patients was *Hira* (Halford et al. 1993a). It is a conserved gene with histone binding properties and is a putative transcription factor. Two homologues in yeast are Hir1p and Hir2p that are transcriptional co-repressors working at the chromatin level in a cell-cycle regulated manner. Several Hira binding proteins have been discovered including Pax3, making it an exciting candidate for branchial arch dysregulation as *Pax3* is expressed in NCCs (Magnaghi et al. 1998). In addition antisense attenuation experiments in the chick cardiac neural crest resulted in an increased frequency of PTA (Farrell et al. 1999). *Hira* may be involved in cell-cycle dependent transcription regulation of a wide range of target genes, possibly acting by altered chromatin structure (Lorain et al. 1998). Recent evidence has shown *Hira* to be involved in senescence associated cell-cycle exit (Zhang et al. 2005a).

Expression studies in mouse and chick have shown *Hira* to be expressed in post gastrulation stages. At e8-9/st6-12 expression is detected in the neuroepithelium, pre- and post migratory neural crest and head mesenchyme. At E9.5/st 18-23 expression is seen in NC derived regions of the head and branchial arches (Roberts et al. 1997), (Wilming et al. 1997).

Targeted mutagenesis in the mouse demonstrated that *Hira* is essential for embryogenesis. Heterozygous embryos were normal, but null embryos on a CD1 genetic background had cardiac looping abnormalities and disturbed cardiac chamber formation (Roberts et al. 2002). Null embryos on a 129S background had gastrulation defects (Roberts et al. 2002). *Hira* is thus an interesting gene in its own right but does not appear to be responsible on its own for the 22q11DS phenotype.

1.4.3.2 UFD1L

Ufd1L gains its name from its homology to the yeast ubiquitin fusion-degradation protein. This is an essential component for yeast protein degradation. *Ufd1L* gathered particular interest as a candidate gene for 22q11DS after a patient was described with a 20kb hemizygous deletion disrupting the *Ufd1L* gene (Yamagishi et al. 1999). In addition, like many other genes within the DGCR, *Ufd1L* is expressed in the limb buds, heart and branchial arches of the developing mouse embryo. Antisense constructs in the chick resulted in an increased incidence of conotruncal septation defects (Yamagishi et al. 2003a). *Ufd1L* has been identified as a downstream target of the *Hand2* transcription factor in *Hand2-/-* embryos (Yamagishi et al. 1999). This is important as *Hand2* mutants have defects of heart development including OFT malformations (Srivastava et al. 1997).

Evidence against *Ufd1L* as a candidate for 22q11DS came from the heterozygous knockout mouse which was normal and the lack of mutations in patients with the phenotype but without the 22q11 deletion (Lindsay et al. 1999a).

1.4.3.3 CRKL

CRKL encodes an SH2-SH3-SH3 adapter protein which is widely expressed and acts within several signalling pathways involved in growth and differentiation (Feller 2001). Mice targeted to lack *Crkl* expression exhibit many features reminiscent of 22q11DS such as craniofacial anomalies, defects within the cranial ganglia, aortic arch defects, outflow tract anomalies, thymus and parathyroid gland anomalies (Guris et al. 2001). However the gene is distal to the DGCR and therefore *CRKL* is not reduced to hemizygosity in all 22q11DS patients and heterozygous mice are normal.

Genes within the DGCR possibly mediating the Psychiatric Phenotype

1.4.3.4 COMT

Catechol-O-methyl transferase (COMT) encodes a protein responsible for the breakdown of dopamine, noradrenaline and adrenaline. Polymorphism within the gene leads to either a high activity (Valine 158) or low activity (Met 18) allele. It has been

postulated that patients with heterozygosity of *COMT* (through a 22q11 deletion) and a Met polymorphism of the other allele may have higher levels of dopamine and catecholamines and therefore more chance of developing schizophrenia. Some studies have supported this association (Lachman et al. 1996), whereas others have disputed it so this relationship is currently ambiguous.

1.4.3.5 PRODH

Proline dehydrogenase (PRODH) is a mitochondrial enzyme that is involved in the transfer of redox potential. Reduced *PRODH* activity caused hyperprolinaemia which is a risk factor for schizoaffective disorder (Jacquet et al. 2004). Proline may be a modulator of neuronal glutaminergic activity. A heterozygous deletion of the entire *PRODH* gene was found in a study of 63 unrelated schizophrenic patients (Jacquet et al. 2002). The same study found loss of function heterozygous mutations in three other patients. This puts the *PRODH* gene high on the list of candidates for a mediator of the psychiatric phenotype seen in 22q11DS patients.

Prodh mutant mice have increased circulating and brain proline levels, and behavioural analysis indicates a decreased prepulse inhibition (PPI) of the startle reflex, suggestive of a defect in sensorineural gating in these mice (Gogos et al. 1999). Decreased PPI is also found in patients with schizophrenia (Parwani et al. 2000).

1.4.3.6 ZDHHC8

ZDHHC8 encodes a putative transmembrane palmitoyltransferase. There is recent evidence using SNP haplotype analysis that ZDHHC8 contributes to the risk of schizophrenia in a sexually dimorphic way. This analysis was prompted by the evidence that knockout mice have a deficit in PPI (Mukai et al. 2004).

1.4.4 Mouse Models

1.4.4.1 Df1 (deficiency 1) mouse

Experiments using mouse models have been used to further dissect the aetiology of 22q11DS. Lindsay *et al* produced a mouse model with a heterozygous deletion (termed *Df1*) within the region in the mouse homologous to the DGCR in humans (Lindsay et al. 1999a). This region is 1.2Mb in length and spans from *Es2* to *Ufd1L* on mouse chromosome 16 establishing hemizygosity of 18 genes (figure 1.4.4.1.1). Although the regions are homologous there is some alteration in gene order in the mouse region when compared to the human deleted region. This is depicted by the coloured lines in figure 1.4.4.1.1. The red lines represent genes without a position change between human and mouse and the blue and green lines represent genes that have changed position. *CLTCL* is present in the human 22q11 region, but a mouse homologue does not exist.

When the *Df1* mice were examined it was found that the homozygous embryos did not implant in the uterus. The heterozygous embryos exhibited a range of cardiac defects similar to those seen in 22q11DS. It was found that 10% of *Df1/+* embryos died at birth. When examined at embryonic day 18 (e18) 26% had cardiac defects including three mice with interrupted aortic arch type B, one with VSD and pulmonary stenosis, six with a retro-oesophageal right subclavian artery and one with aberrant origin of the right subclavian artery. At term 18% had defects including one with a right sided aortic arch and the remainder with retro-oesophageal subclavian artery. In addition it was shown that *Df1* haploinsufficiency was the cause, as the phenotype was rescued when the *Df1/+* mice were bred with mice carrying a duplication of the *Df1* genes (Lindsay et al. 1999a).

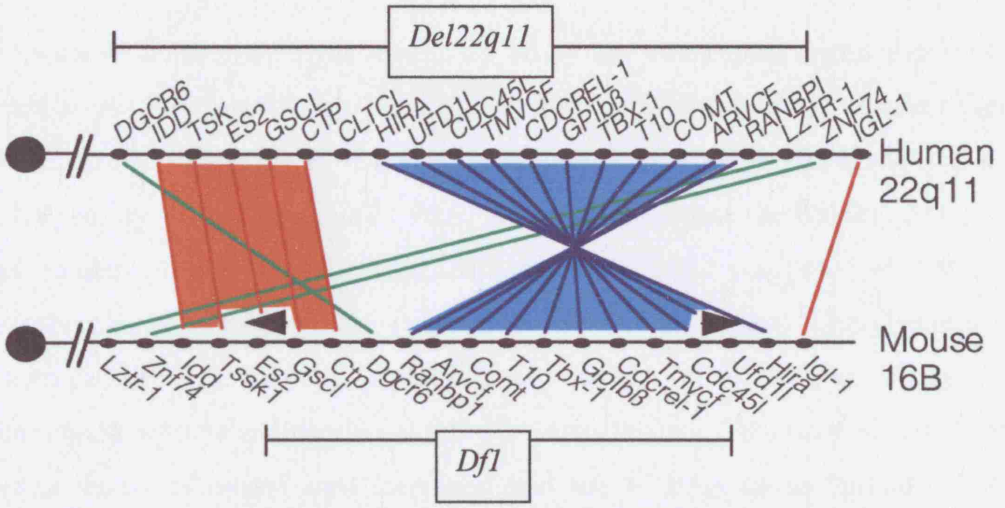


Figure 1.4.4.1.1 Genes deleted in the Df1 mouse
Gene content and order in the region commonly deleted in 22q11DS patients, compared with the syntenic area of mouse chromosome 16 deleted in *Df1* mice. Figure taken from Lindsay *et al* 2001.

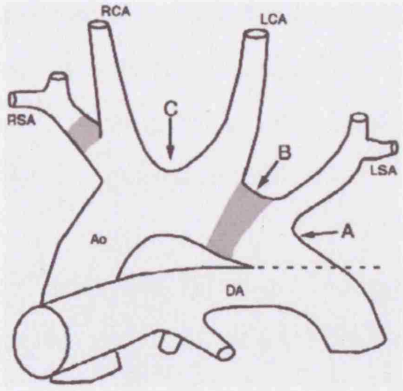


Figure 1.4.4.1.2a Diagram of the aortic arch late in foetal life.
The shaded areas in the aortic arch and proximal right subclavian artery represent the segments that derive from the 4th arch arteries.
Ao.aorta; DA: ductus arteriosus;
LCA: left carotid artery; LSA: left subclavian artey; RCA: right carotid artery; RSA; right sublcavian artery. A,B and C show the 3 different sites of aortic arch interruption.

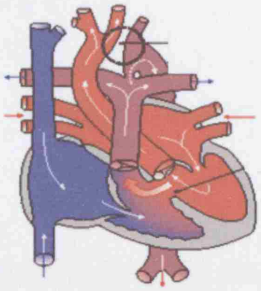


Figure 1.4.4.1.2b Interrupted aortic arch type B.

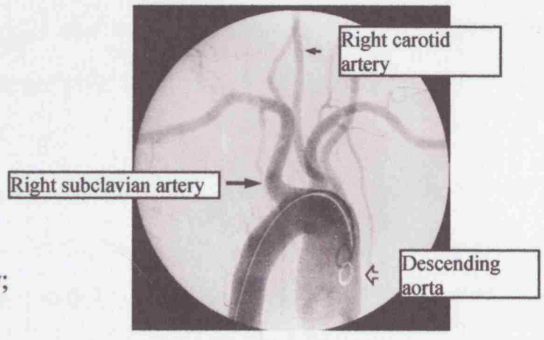


Figure 1.4.4.1.2c Aberrant right subclavian artery seen at a cardiac angiogram

Figure 1.4.4.1.2b and c. Examples of typical abnormalities seen in patients with 22q11DS and *Df1* mice, thought to originate from 4th arch artery defects.

Defects such as an aberrant right subclavian artery and interrupted aortic arch type B suggested an aetiology involving the 4th pharyngeal arch artery (4th PAA) (see figure 1.4.4.1.2). Indeed, when the 4th PAAs were examined more closely at E10.5 100% of the *Df1/+* embryos had abnormally small 4th PAAs (Lindsay & Baldini 2001). A reduced number of embryos with aortic arch defects at e18.5 compared with at term, suggests there must be some genetic rescue from the 4th PAA defects. This phenomenon was examined more closely in Lindsay *et al* (Lindsay et al. 1999b). Markers for vascular smooth muscle staining in the arteries were deficient. This suggested that there was either failure of neural crest migration into the 4th PAA or an inability of post migratory neural crest differentiation into vascular smooth muscle due to a lack of response to developmental signals. By looking at neural crest staining with *Crabp1* the authors showed that neural crest migration does occur (Lindsay et al. 1999b).

The authors Lindsay and Baldini postulated that the 4th PAA defects are likely to result from a local signalling pathway defect between endothelial cells or pharyngeal endoderm and neural crest derived mesenchyme leading to a lack of differentiation into vascular smooth muscle cells (Lindsay & Baldini 2001). This hypothesis is favoured over a more general problem with angiogenesis as the other arch arteries are rarely affected in the *Df1* animal model.

Furthermore Taddei and Lindsay (Taddei et al. 2001) showed that the penetrance of the heart defects was increased on a C57Bl/6 mouse background (compared to 129SvEdBrd x C57Bl/6) with 50% of heterozygous mice showing defects. In addition the mice showed varying degrees of thymic hypoplasia and anomalously positioned parathyroid glands, modelling even more closely the phenotypic features of 22q11DS.

Similar experiments created a larger (1.5Mb) deletion termed *Lgdel* (Merscher et al. 2001). The phenotype of *Lgdel/+* mice was similar to that seen in *Df1/+* although the perinatal mortality rate was higher. Figure 1.4.4.1.3 illustrates the location of the *Df1* and *Lgdel* deletion constructs made.

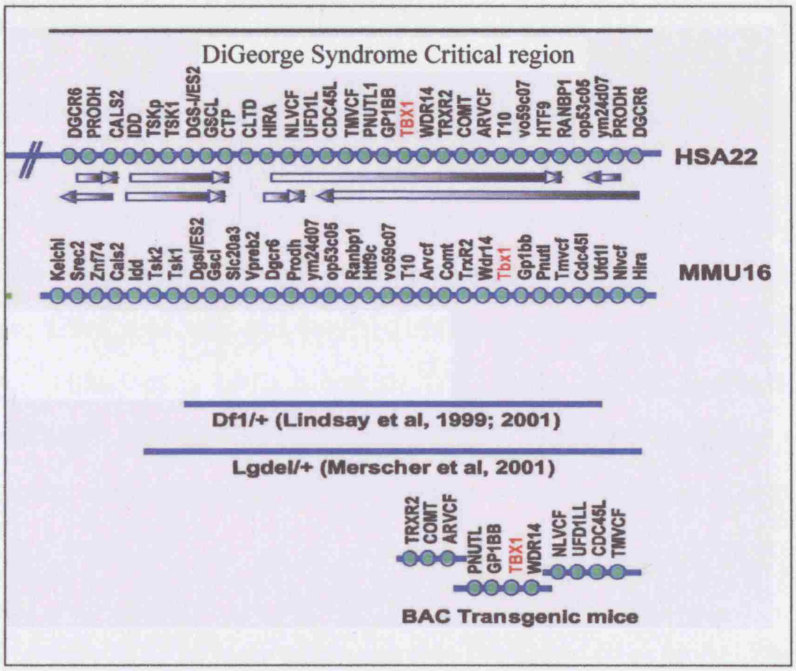


Figure 1.4.4.1.3 22q11DS mouse deletion constructs.

Detailed diagram of the *Df1* and the human 22q11.2 region. Underneath are examples of two of the mouse deletion constructs made. The BAC transgenic constructs used to rescue the phenotype are also shown. *Tbx1* is highlighted in red. Figure adapted from Merscher *et al*. 2001.

1.4.4.1.1 Behavioural Phenotype of the Df1 mouse

Interestingly the *Df1* mouse shows defects in its prepulse inhibition (PPI) to a startle response and in Pavlovian conditioned fear tests (Paylor et al. 2001). Defects in PPI are also seen in patients with schizophrenia and reflect their inability to ignore extraneous sounds and concentrate on the task in hand. The conditioned fear test results are interesting as they signify defects in learning and memory. In the experiments by Paylor *et al* the mice showed particular difficulty with complex tasks involving memory over 24 hours whereas short term memory tasks test after 1 hour were intact (Paylor et al. 2001). This is similar to the problems seen in children with 22q11DS where rote learning and short term memory is intact, but the learning of more complex tasks is sometimes impaired (Woodin et al. 2001). The *Df1* mouse provides a useful model for the psychiatric and behavioural anomalies seen in patients with 22q11DS.

1.4.4.2 The Tbx1 Mouse Model

Once it was established that the *Df1* mouse replicated some of the phenotypic features of 22q11DS, nested deletions and duplications in the mouse region were produced to try and establish if one gene in particular was responsible for the phenotype. A critical region for the defects seen in *Df1/+* mice was defined through genetic complementation between different sized deletions and bacterial artificial chromosome (BAC)-containing transgenic mice (Lindsay et al. 2001; Merscher et al. 2001). A BAC that harboured four human genes-*GP1Bβ, PNUTL1, TBX1* and *WDR14* provided complete phenotypic rescue in most mice, demonstrating that one or more of the four genes is responsible for the defect (figure 1.4.4.1.3).

One of the four genes *TBX1* is highly expressed in the pharyngeal arches during mouse embryonic development (Chapman et al. 1996). This gene was targeted for inactivation by three different research groups (Jerome & Papaioannou 2001; Lindsay et al. 2001; Merscher et al. 2001). *Tbx1* heterozygous mice had mild cardiovascular defects similar to those seen in *Df1* mice, but did not show reduced viability. Homozygous mice had more severe defects and died perinatally. These defects are summarised in table 1.4.4.2 and illustrated in figure 1.4.4.2. The phenotype for 22q11DS has been reproduced in its severest form in the *Tbx1* homozygous mouse. The mice have persistent truncus arteriosus, cleft palates, absent thymus and parathyroid glands and craniofacial features similar to 22q11DS patients. A histological analysis of embryos between E10.5 and E12.5 revealed a severely narrowed and laterally constricted pharynx. By examining earlier embryos it was seen that the 1st pharyngeal arch formed correctly, but the second pharyngeal arch was hypoplastic and arches 3-6 were not present at all, accounting for the abnormalities seen in the structure that derive from these arches (Jerome & Papaioannou 2001; Lindsay et al. 2001; Merscher et al. 2001).

Table 1.4.4.2 Summary of the abnormalities found in the *Tbx1* deficient mouse *(references in text)*

Abnormality	*Tbx1* Heterozygote	*Tbx1* Homozygote
e7.5	No abnormalities seen.	No abnormalities seen.
E8.5		Only 1st arch and pouch visible.
E9.5		1st arch misshapen, 2nd arch hypoplastic, 3rd arch and pouch missing in 90% of embryos.
E10.5		2nd arch hypoplastic, no 3-6th arches
E11.5	Decreased or absent 4th arch arteries in 45%.	Decreased or absent 4th arch arteries
Viability	Viable.	Die at birth, Oedematous.
Cardiovascular		
Pharyngeal arch arteries and outflow tract phenotype.	50% have abnormalities including aberrant right subclavian, retro-oesophageal right subclavian artery, IAA, and cervical aortic arch.	Persistent Truncus arteriosus (PTA), PTA plus RAA, PTA plus double aortic arch
Heart		Small ventricles.
Craniofacial		
Palate	No abnormalities seen.	Cleft.
Middle ear	No abnormalities seen.	Ossicles abnormal/ semicircular canals abnormal.
Larynx	No abnormalities seen.	Hyoid/thyroid/cricoid cartilages abnormal.
Mandible	No abnormalities seen.	Micrognathic. No coronoid process.
Teeth	No abnormalities seen.	Upper incisors missing/malformed.
Cervical spine	No abnormalities seen.	Short neck, atlas lacking anterior arch.
Cranial bones	No abnormalities seen.	Basioccipital/Basisphenoid bones fused.
Facial bones	No abnormalities seen.	Abnormal zygomatic arch/ temporal bone/ tympanic ring.
Glandular structures		
Parathyroid gland	No abnormalities seen.	Absent.
Thymus	No abnormalities seen.	Absent/single lobe.
T cells	No abnormalities seen.	No abnormalities seen.
Thyroid gland	No abnormalities seen.	Present.
Salivary Glands	No abnormalities seen.	Missing sublingual/submaxillary salivary glands.
Ears	No abnormalities seen.	Inner ear, semicircular canals and pinnae abnormal.

Figure 1.4.4.2. Abnormalities in the *Tbx1* mouse

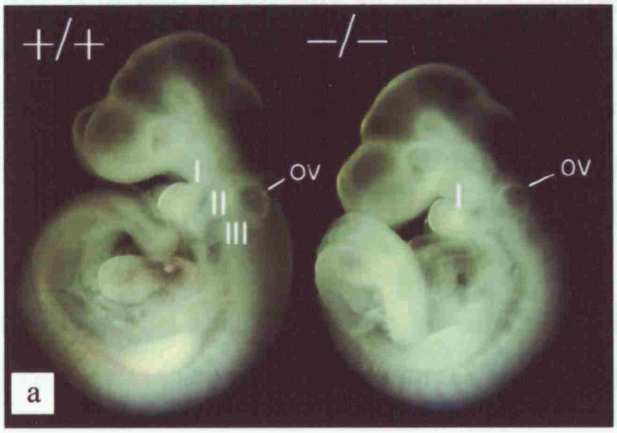


Figure 1.4.4.2a Whole mounts of mouse wild type (+/+) and *Tbx1* homozygous embryos (-/-) at e9.5.
Note pharyngeal arch's I, II and III in wild type, but in the homozygote pharyngeal arch II is hypoplastic and III is missing.

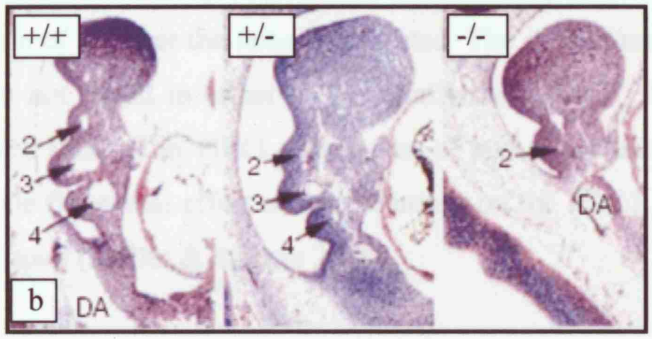


Figure 1.4.4.2b The pharyngeal arch phenotype in *Tbx1* e11 embryos
Parasagittal histological sections through the pharyngeal arch region in wild type, *Tbx1 +/- and Tbx1-/-* embryos. 2,3 and 4 denote the pharyngeal arch arteries. Note the hypoplasia of the 4th PAA in the *Tbx1* +/- embryos (identical to that seen in *Df1*/+ embryos). In the *Tbx1* -/- embryo the 2nd PAA is hypoplastic and the 3rd and 4th PAA's are absent. (Lindsay et al. 2001). DA: dorsal aorta.

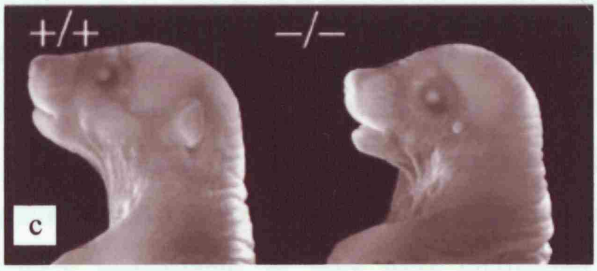


Figure 1.4.4.2c Tbx1 wild-type and homozygous mutant neonates.
Comparison of the external appearance of *Tbx1* +/+ and *Tbx1* -/- mice. Note the abnormal pinna and the micrognathia.
(Jerome and Papaioannou 2001).

1.4.5 TBX1 mutations have been found in patients with 22q11DS

TBX1 is deleted in virtually all 22q11DS patients (Chieffo et al. 1997), but it is only recently that mutations in *TBX1* been found in patients not deleted at 22q11.2 (Paylor et al. 2006; Gong et al. 2001; Yagi et al. 2003; Stoller & Epstein 2005). Yagi *et al* (Yagi et al. 2003) screened the DNA of 10 non-deleted patients with the same phenotype as 22q11DS and found three patients with mutations in *TBX1*. These mutations are illustrated in figure 1.4.5. Although these patients had most of the features of 22q11DS they did not have significant learning difficulties. This would be consistent with the lack of *Tbx1* expression in the mouse brain (Maynard et al. 2003).

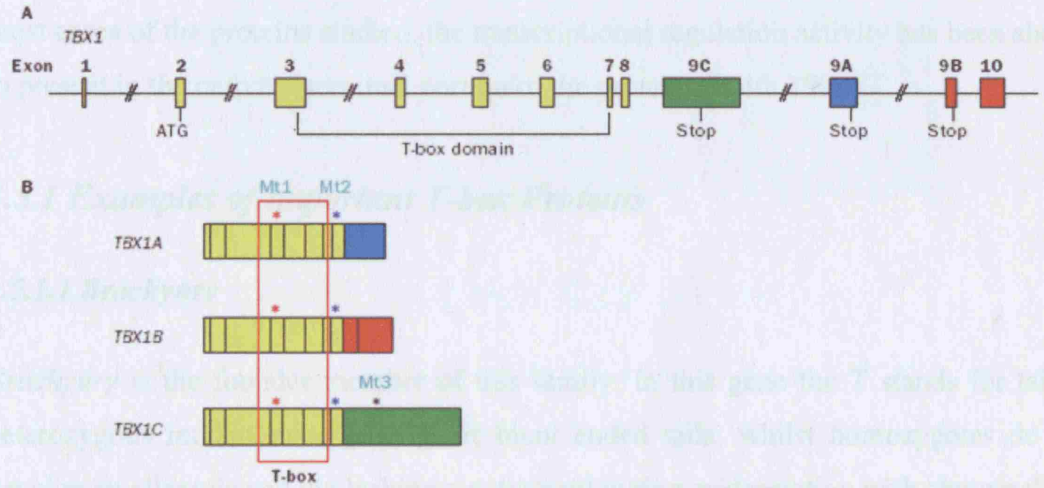
Two out of three of the mutations in *TBX1* were missense mutations; one in the T-box (DNA binding domain) and one in the C terminal (activation domain). The first mutation was F148Y (443T→A, Mt1) and was not found to be present in the patient's mother with no mention of whether the father was tested. The second mutation, G310S (928G→A, Mt2), was not found in either of the unaffected parents. Both mutations were found in conserved residues in TBX1 orthologues of human, mouse, Xenopus and zebrafish. At present the functional effect of the mutations on the TBX1 protein is only beginning to be understood (Stoller & Epstein 2005).

The third mutation, 1223delC (Mt3), resulted in a frameshift leading to a stop codon after 51 codons. This mutation in exon9c would result in a truncated TBX1 protein without its C terminus. Exon 9c is predicted to be part of only one of three putative isoforms of the TBX1 protein (isoform C, the others being A and B). The C isoform is thought to be the most important as it is the most abundant and is fully conserved during evolution (Chieffo et al. 1997). This same mutation was found in 3 affected family members, with variable phenotypic expression.

It is intriguing as to why so few human mutations have been found. It may be that there is an ascertainment bias in that most groups have been looking at patients with learning difficulties as part of the phenotype. Gong *et al* screened 105 non deleted patients and found eight common polymorphisms and 10 rare variants (Gong et al. 2001). Unfortunately with the rare variants, either there were no amino acid changes, or the parents were unavailable, or they were unaffected carriers. There is however, no

mention by the authors of the clinical investigations done in the parents and the phenotype is known to be extremely variable in 22q11DS. In conclusion it seems that *TBX1* mutations in non deleted patients appear to be rare and further laboratory work is needed to assess the significance of the mutations found.

Figure 1.4.5 Structural and functional organisation of *Tbx1*



a) Schematic representation of the alternatively processed transcripts, *TBX1A*, *TBX1B*, *TBX1C*.

b) gene structure and site of human mutations 443T→A (Mt1), 928G→A (Mt2), and 1223delC (Mt3). The T-box domain is represented by an open red square and spans exons 3-7. * The 3 mutations in *TBX1* found by Yagi et al (Red F148Y; blue G310S; Black 1223delC) (Yagi et al. 2003).

1.5 What is known about T-box genes?

It is fairly clear that *Tbx1* is the main gene responsible for many of the physical features of 22q11DS. What is known about this family of transcription factors? T-box transcription factors comprise an ancient gene family that contributes to the differentiation and organogenesis of many systems in vertebrates and invertebrates.

The T-box is defined as the minimal region within the T-box protein that is both necessary and sufficient for sequence specific DNA binding. Despite the sequence

variations within the T-box between family members, examination of downstream targets and binding site selection experiments show that all members of the family so far examined bind to the DNA consensus sequence TCACACCT (Conlon et al. 1996; Ataliotis et al. 2005 ; Ghosh et al. 2001). Some T-box family members preferentially bind to two or more duplicates of this sequence arranged in various orientations (Smith 1999).

T-box proteins have been demonstrated to function both as transcriptional activators {e.g. Tbx5 {Horb et al. 1999}} and as repressors {e.g. Tbx2 {Carreira et al. 1998}}. In most cases of the proteins studied, the transcriptional regulation activity has been shown to present in the carboxy-terminal portion of the protein (Smith 1999).

1.5.1 Examples of important T-box Proteins

1.5.1.1 Brachyury

Brachyury is the founder member of this family. In this gene the T stands for tail as heterozygous mutant mice have short blunt ended tails, whilst homozygotes do not develop an allantois and die lacking a notochord during midgestation with abnormalities in mesoderm derived tissues (Wilkinson et al. 1990; Chelsey 1935). Crystallographic analysis of the Xenopus homolog of *brachyury, Xbra* revealed it to bind as a dimer to the palindromic version of the T-box consensus sequence T\{G/C}ACACCTAGGTGTGAAATT, with the C-terminal helix being embedded into an enlarged minor groove of DNA (figure 1.5.1).

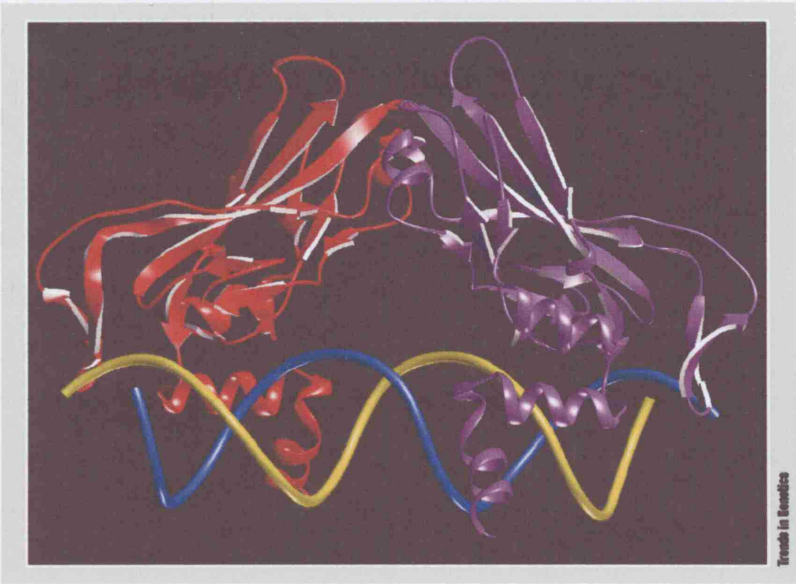


Figure 1.5.1. Structure of a typical T-box protein

Ribbon diagram of two *Xbra* T-box domains bound to the palindromic DNA motif as a dimer. Red/Purple ribbons: two T-box domains Blue/Yellow: two DNA strands (Smith 1999).

1.5.1.2 TBX3: Ulnar-Mammary Syndrome

Table 1.5.1 is a summary of the existing T-box genes and their currently known function. Ulnar mammary syndrome (UMS, OMIM 181450) is caused by mutations in *TBX3* (Bamshad et al. 1997). It affects the ulnar ray of the limb producing defects such as a missing fifth finger/ hypoplasia/ absent ulnar. In addition patients often have hypoplasia of the breast tissue and can have genital and tooth abnormalities. There is no known genotype-phenotype correlation in the syndrome, suggesting that like many T-box genes *Tbx3* acts by a mechanism of haploinsufficiency. Breasts, tooth and genital development all rely on inductive interaction between epithelial tissue and underlying mesenchyme and *TBX3* may have a common role in these developmental processes.

1.5.1.3 TBX5: Holt-Oram syndrome

Holt-Oram Syndrome (HOS, OMIM 142900) is characterised by defects in radial ray bones such as hypoplastic thumbs or malformed radii. Heart defects such as atrial septal

defects (ASD)s also occur, usually associated with conduction defects and ECG abnormalities. Mutations in *TBX5* have been identified in both familial and sporadic cases of HOS (Li et al. 1997b). The majority of mutations cause haploinsufficiency of *TBX5*(Li et al. 1997b). There is some phenotype-genotype correlation. Skeletal defects are more prevalent in patients with mutations in the C terminal end of the T-box which contacts the minor groove of the DNA. Aberrant heart development is predominantly seen in patients with missense mutations in the N terminal end of the T-box that contacts the major groove (Basson et al. 1997).

The cardiac and limb defects are reproduced in a mouse model with a heterozygous mutation in *Tbx5* (Bruneau et al. 2001). In similarity to *Tbx1*, the genetic background of the mice significantly affects the phenotype suggesting that modifying genes are important. Mouse studies have identified downstream targets of *Tbx5* to include *ANF, Connexin 40, Mlc2V, Irx4, Hey2, Nkx2.5 and Gata4* (Stennard & Harvey 2005). In addition mutations in *Nkx2.5* in humans can produce atrial and tricuspid valve abnormalities as well as DORV and TOF (Tanaka et al. 1999).

1.5.1.4 Tbx19/ Isolated ACTH deficiency

Tbx19, also known as *Tpit*, was identified as being required for the expression of pro-opiomelanocortin (POMC) in the cortotroph and melanotroph pituitary lineages (Lamolet et al. 2001). In humans the absence of POMC leads to lack of adrenal cortical releasing hormone (ACTH) resulting in adrenal insufficiency (OMIM 201400). Mutations in *TBX19* were found in patients with this condition (Lamolet et al. 2001). Inheritance is recessive and mutations are loss-of function alleles.

1.5.1.5 TBX22/Cleft Palate with Ankyloglossia

Cleft Palate with Ankyloglossia (CPX, OMIM 303400) is an X linked dominant condition mapping to Xp21. Mutations in *TBX22* have been found in patients with CPX (Braybrook et al. 2001). *TGFβ2* is a possible target for *TBX22* as homozygous deletions in the mouse produce cleft palates. Thus *TBX22* may act in a pathway directing the transformation of epithelium into mesenchyme during formation of a continuous palate.

Table 1.5.1 showing existing T-box genes and their known function

T-box member	Function	Human Disease	Mouse model	References
Tbx1	Transcriptional activator involved in pharyngeal arch development.	Mutations found in 22q11DS.	PTA. Cleft/thymus/parathyroid/ craniofacial abnormalities. See table 1.4.4.2 and text.	(Lindsay et al. 2001; Jerome & Papaioannou 2001; Merscher et al. 2001; Yagi et al. 2003)
Tbx2	Transcriptional repressor. Represses melanocyte specific TRP-1 promoter.	Implicated in breast and ovarian cancer.	Homozygous mouse: AVSD, hindlimb abnormalities and digit duplication	(Harrelson et al. 2004)
Tbx3	Acts as a transcriptional repressor. Binds as a monomer. Impinges on p53 pathway to suppress apoptosis and facilitate call transformation.	Ulnar Mammary Syndrome.	Homozygous mouse has hindlimb and forelimb defects, mammary gland and yolk sac defects.	(Bamshad et al. 1997)
Tbx4	Involved in hindlimb specification. Downstream target of *Pitx1*.	Unknown.	Homozygous mice fail to undergo chorioallantoic fusion and die at E10.5. They show malformations in limb bud development.	{Naiche et al. 2003 }
Tbx5	Transcriptional activator. *BMP4* is upstream of *Tbx5*.	Holt Oram Syndrome.	*Tbx5* heterozygous mice: radial ray deformities/ cardiac: ASD/ conduction defects.	(Bruneau et al. 2001)
Tbx6	Involved in somite and muscle differentiation.	Candidate for spondylocostal dysostosis.	Homozygous mouse has 3 neural tubes. Tbx6 is required for somite patterning.	(Chapman & Papaioannou 1998; White et al. 2003)
Tbx7	Unknown. Found in *C.elegans*. Redundant name for *Tbx10.*	Unknown		(Agulnik et al. 1997)
Tbx8	Unknown. Found in *C.elegans*-contributes to formation of hypodermis and body wall muscles. Redundant name for *Tbx15.*	Unknown		(Agulnik et al. 1997; Andachi 2004)
Tbx9	Found in *C.elegans*. Deletion in *C.elegans* showed a partial loss of spatial abnormality in body-wall muscles.	Unknown		(Agulnik et al. 1997)
Tbx10	Closely related to *Tbx1*.	Unknown	Gain of function in the dancer mouse result in cleft lip and palate.	(Law et al. 1998) {Bush et al. 2004}

T-box member	Function	Human Disease	Mouse model	References
Tbx11	Unknown. Found in *C.elegans.*	Unknown		(Agulnik et al. 1997)
Tbx12	Unknown. Has a short C terminus and expressed in retina and heart. *Tbx1* subfamily. Activation and repression domains are in N terminus.	Unknown		(Carson et al. 2000)
Tbx13	Redundant name for *Tbx10*			
Tbx14	Redundant name for *Tbx15*			
Tbx15	Same subfamily as *Tbx1*. Expressed strongly in craniofacial region in mouse.	Candidate gene for acromegaloid facial syndrome	Responsible for dorsoventral coat patterning in the droop ear mouse.	(Agulnik et al. 1998; Candille et al. 2004)
Tbx16	Expression pattern defined in Xenopus. Gives rise to the lateral and ventral mesoderm during gastrulation.	Unknown		(Ruvinsky et al. 1998)
Tbx17	Unknown. Found in *C.elegans.*	Unknown		(Agulnik et al. 1997)
Tbx18	Unknown. Important for somite formation in the chick.	Unknown		(Tanaka & Tickle 2004)
Tbx19	*Tpit* activates *POMC* cell differentiation (see text)	ACTH deficiency	*Tpit* homozygous mice have adrenal insufficiency due to lack of corticotroph differentiation and POMC expression.	(Yi et al. 1999)
Tbx20	Expressed in cardiac tissue. C terminus can act as a transcriptional repressor at the *ANF* promoter. Interacts with *Nkx2.5, Gata4* and *Gata5* in cardiac tissue.	Unknown		(Plageman, Jr. & Yutzey 2004; Stennard et al. 2003)
Tbx21	Th1 lineage, Lung and spleen (adult).		Th1 lineage commitment in immune system development.	(Faedo et al. 2002)
Tbx22	Expressed in palatal shelves and the frenulum of the tongue. Function largely unknown. T-box region 20 amino acids shorter than other T-box domains.	X-linked cleft palate and ankyloglossia.		(Braybrook et al. 2001)

1.6 Tbx1

The *Tbx1* -/- mouse has virtually all of the features of 22q11DS, but in heterozygous mice only a cardiac phenotype is apparent (Lindsay et al. 2001; Jerome & Papaioannou 2001; Merscher et al. 2001). This may indicate a difference in dosage sensitivity between mice and humans, and other genes within the DGCR may be involved in creating the phenotype in humans.

Work from the Molecular Medicine Unit, ICH has shown that Xenopus *Tbx1* is a transcriptional activator (Ataliotis et al. 2005) and that it binds the same palindromic nucleotide sequence as brachyury (Ataliotis-personal communication). Work is underway to establish if the published human mutations in *TBX1* have any functional activity at this site (Paylor et al. 2006; Stoller & Epstein 2005)

1.6.1 Tbx1 Expression Domains

The expression pattern of *Tbx1* in mouse was first described by Chapman in 1996 (Chapman et al. 1996). Fate mapping by Brown *et al* showed that *Tbx1* is expressed early on in mesodermal derivatives such as the lateral and intermediate plate mesoderm including pre-cursors of cardiac and smooth muscle (Brown et al. 2004). As the pharyngeal arches are forming between e9 and e11 in the mouse, *Tbx1* is expressed in the pharyngeal mesoderm, pharyngeal endoderm and head mesenchyme. *Tbx1* is also expressed in the epithelium of the otic vesicle in the epithelium of the endodermal pouches. Endodermal expression in the pharyngeal pouches is temporally regulated such that expression is first noted in the first pouch, followed by sequential expression in more caudal pouches (Vitelli et al. 2002b). Expression is also observed in the sclerotome consistent with the skeletal defects seen in *Tbx1* null mice (Jerome & Papaioannou 2001). *Tbx1* has more recently been found to be expressed in the outflow tract of the heart (Vitelli et al. 2002a). The patterns of *Tbx1* expression are shown in figure 1.6.1.

Figure 1.6.1 Mouse *Tbx1* Expression Pattern

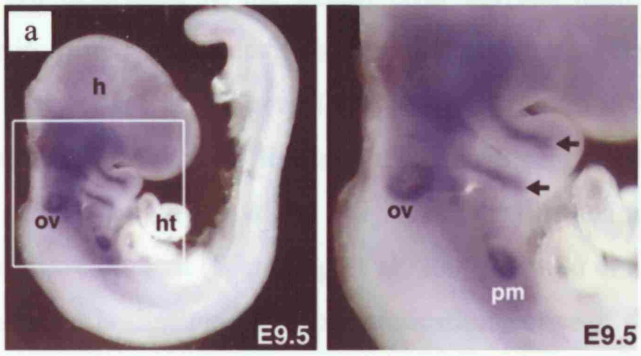


Figure 1.6.1a Whole mount *in situ* hybridisation of *Tbx1* in wild type e9.5 mouse embryo.
Expression is shown in the head (h), otic vesicle (OV) and pharyngeal arches. Higher magnification shown on the right (the boxed area magnified) revealed that expression was localised to the mesodermal core of the pharyngeal arches (arrows). ht: heart, pm: periaortic mesenchyme. Figure from Garg et al. 2001.

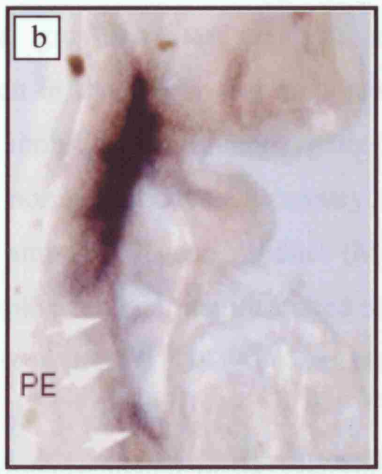


Figure 1.6.1b Whole mount *in situ* hybridisation of *Tbx1* in e9.5 mouse embryo.

Arrowheads indicate expression in the pharyngeal endoderm (PE). Figure from Lindsay et al. 2001.

Figure 1.6.1c Expression pattern of *Tbx1* in e10.5 mouse embryo.

Sagittal section of e10.5 embryo through the pharyngeal arches reveals *Tbx1* expression surrounding the aortic arch arteries (numbered 2, 3 and 4), especially around artery 4. Br: brain, Ht: heart. Figure from Merscher et al. 2001.

Tbx1 transcripts are not detectable in the neural crest derived mesenchyme of the pharyngeal arches nor in primary cultures of chick neural crest cells (Garg et al. 2001). These results suggest that defects of the neural crest derived tissues in 22q11DS could occur in a non-cell autonomous fashion. Recent studies have suggested that defects in pharyngeal epithelial endoderm might affect migration or differentiation of NCCs during pharyngeal arch development (Wendling et al. 2000; Arnold et al. 2006). Focus has now turned to the signals that mediate communication between the neural crest, endodermal and mesodermal cells to uncover the molecular pathways involving *Tbx1* (figure 1.6.4).

1.6.2 Tbx1 in aortic arch patterning

It has recently been established that there is genetic link between Fibroblast growth factor 8 (*Fgf8)* and *Tbx1* (Abu-Issa et al. 2002; Frank et al. 2002; Vitelli et al. 2002b). Fibroblast growth factor 8 (*Fgf8*) is a secreted signalling molecule with dynamic expression in the pharyngeal arch ectoderm and endoderm, and the developing brain, face and limbs. As *Fgf8* homozygous mice die of gastrulation defects and heterozygous mice are normal, it became necessary to analyse a hypomorphic allele which produces a reduced amount of gene product (Meyers et al. 1998). Analysis of mouse mutants hypomorphic for *Fgf8* demonstrated similar cardiovascular and craniofacial phenotypes to those seen in 22q11DS (Abu-Issa et al. 2002; Frank et al. 2002).

Fgf8 is downregulated in the pharyngeal endoderm of *Tbx1* mutants, and mice trans-heterozygous for both *Tbx1* and *Fgf8* have a higher penetrance of aortic arch defects than *Tbx1* heterozygotes (Vitelli et al. 2002). It is already well established that during embryogenesis epithelially produced *Fgf8* provides survival, differentiation and patterning signals to adjacent mesenchyme (Tucker et al. 1999).

Tbx1 causes growth/remodelling defects of the 4th PAAs in a cell non-autonomous way (as *Tbx1* is not expressed in the structural components of the artery, but in the endoderm). Vitelli *et al* showed that these defects are due to abnormal migration of NCCs (Vitelli et al. 2002a). As *Fgf8* appears to be responsible for NCC survival it seems that *Fgf8* must provide a modifying effect on NCC migration and survival.

1.6.3 Tbx1 in Outflow Tract formation

Prior to E11.5, there is a common OFT, the truncus arteriosus, which connects the heart to the systemic circulation. Subsequently, the OFT undergoes septation to separate aortic and pulmonary flows. *Tbx1* is expressed in a population of myocardial precursor cells known as the secondary or anterior heart field (AHF), which originate from the splanchnic mesoderm (caudal to the pharynx and anterior to the heart tube) and go on to form the myocardium of the OFT. Neural crest derived cells are also required for OFT septation and make up most of the aorto-pulmonary septum.

100% of mice homozygous for the null mutation in *Tbx1* have a failure of OFT septation leading to PTA, suggesting a crucial role for *Tbx1* in OFT development (Lindsay et al. 2001; Jerome & Papaioannou 2001; Merscher et al. 2001). This has been shown to be independent of its role in pharyngeal arch patterning (Xu et al. 2004).

It has been demonstrated *Fgf10* expression is decreased in the anterior heart field cells of *Tbx1* nulls (Kochilas et al. 2002; Vitelli et al. 2002b) and *in vitro Fgf10* is a direct transcriptional target of *Tbx1* (Xu et al. 2004). However, mice homozygous for a null mutation in *Fgf10* do not have OFT defects (Sekine et al. 1999) suggesting that *Fgf10* alone is not responsible for medicating OFT defects in *Tbx1* null mice. It has recently been demonstrated that *Fgf8* expression is also decreased in the AHF cells of *Tbx1* hypomorphs (Hu et al. 2004) and tissue specific deletion of *Fgf8* in *Tbx1* specific domains results in cardiac OFT defects (Brown et al. 2004). Thus both *Fgf8* and *Fgf10* are important for OFT formation directed by *Tbx1*, probably in a cell non-autonomous fashion (Xu et al. 2004).

1.6.4 Molecular regulation and modification of the 22q11DS phenotype

It is known the 22q11DS phenotype is highly variable, even between members of the same family. In addition it has been shown that changing the background in *Df1/+* mice can alter the penetrance of the phenotype (Taddei et al. 2001). It is therefore likely that modifier genes exist (figure 1.6.4).

As the mice deficient in *Tbx1* were first described in 2001, potential interactors up and downstream of *Tbx1* have only recently come to light. *Vegf* (Stalmans et al. 2003) and

RaldH2 (Vermot et al. 2003) deficient mice replicate the phenotype of 22q11DS and are therefore likely to be involved in similar pathways to *Tbx1*.

1.6.4.1 VEGF may act as a genetic modifier in 22q11DS

A recent study of the *v*ascular *e*ndothelial *g*rowth *f*actor (VEGF) gene has provided evidence for a genetic modifier in 22q11DS (Stalmans et al. 2003). *Vegf* has an essential role in vasculogenesis and angiogenesis (Gerber et al. 1999; Peters et al. 1993). In the mouse *Vegf* is expressed at e10 in the endoderm of the 4th PAA, the aortic sac, the outflow tract, the maxillary and mandibular prominences, the palatal midline and the thymus.

At least three isoforms of *Vegf* exist; $Vegf^{120}$, $Vefg^{164}$, and $Vegf^{188}$ (the number of residues is given in superscript). The role of these isoforms has been studied by generating knock-in mice, expressing a single isoform (Stalmans et al. 2003).

Mice expressing only $Vefg^{164/164}$ are normal. However mice expressing the $Vegf^{120/120}$ and $Vegf^{188/188}$ variants died in the neonatal period due to cardiac malformations similar to those seen in 22q11DS such as IAA-B, right aberrant subclavian artery and hypoplasia of the pulmonary trunk (Stalmans et al. 2003). In addition the embryos of the $Vegf^{120/120}$ variant had craniofacial defects such as cleft palate and micrognathia and thymic and parathyroid anomalies (Stalmans et al. 2003).

Further studies with RTQPCR and *in situ* hybridisation showed that *Tbx1* expression was decreased in $Vegf^{120/120}$ mutants. In addition Stalmans *et al* used morpholino oligomer injection to test for genetic interactions between *Tbx1* and *Vegf* in the zebrafish. *Vegf* knockdown increased the penetrance of PAA defects in a *Tbx1* knock-down model in a dose dependent manner (Stalmans et al. 2003).

The defects observed in the mice were not attributed to NCC abnormalities, but were thought to be due to a decreased vascular density of structures such as the mandible and the thymus. It was postulated that a decreased vascular perfusion of these structures resulted in the anomalies noted, a theory that has been extended to account for some of the defects seen in patients with 22q11DS such as the genital, renal and hand

abnormalities. The evidence so far extends the idea of genetic modifiers playing an important role in 22q11DS.

1.6.4.2 Sonic hedgehog regulates Tbx1 expression

Sonic hedgehog (*Shh*) encodes a signalling protein that is instrumental in patterning the early embryo. *Shh* mutant mice die soon after birth with severe craniofacial defects (Chiang et al. 1996). They fail to maintain *Tbx1* expression and chick embryos exposed to *Shh* soaked beads responded by up-regulating *Tbx1* expression in the pharyngeal arches (Garg et al. 2001).

Yamagishi *et al* (Yamagishi et al. 2003b) demonstrated a *cis*-acting genomic regulatory region –14.3 to –13.2 kb upstream of *Tbx1* that controls its expression and is responsive to *Shh* signalling. The element is shown to be a binding site for winged helix/forkhead box (Fox) containing transcription factors and is essential for regulation of *Tbx1* transcription of the pharyngeal endoderm and head mesenchyme (Yamagishi et al. 2003b). *Foxa2* is required for endoderm development and *Foxc1* and *Foxc2* are required for head mesenchyme and aortic arch formation {Iida et al. 1997}.

Mice lacking any two of the four alleles of *Foxc1* or *Foxc2* display aortic arch malformations (Winnier et al. 1995). This suggests that FOXC proteins and TBX1 function in a common pathway. In addition approximately 15% of patients with *FOXC2* mutations, a cause for lymphoedema-distichiasis syndrome, have conotruncal cardiac defects {Fang et al. 2000}.

Yamagishi *et al* propose that *Tbx1* is a direct transcriptional target of Fox proteins and that *Foxa2* and *Foxc2* may serve an intermediary role in *Shh* regulation of *Tbx1* (Yamagishi et al. 2003b).

1.6.4.3 The Role of Retinoic Acid

It is well documented that antenatal exposure to vitamin A, a precursor for retinoic acid (RA) can result in defects seen in 22q11DS. Isoretinoin which has been used as a treatment for acne results in congenital malformations similar to those seen in 22q11DS (Lammer et al. 1986). Furthermore administration of RA to rodent embryos also

produces a 22q11DS phenotype with craniofacial, cardiac outflow tract and thymic abnormalities (Mulder et al. 1998; Mulder et al. 2000). *Fgf8* expression is regulated by RA (Niederreither et al. 2001; Vermot et al. 2003) and RA is responsible to maintaining *Fgf8* and *Shh* expression (Schneider et al. 2001).

A number of studies indicate a role for retinoids in pharyngeal patterning (Lohnes et al. 1994; Wendling et al. 2001). A mouse hypomorphic for the retinoic acid synthesis enzyme *Raldh2* showed features consistent with *Tbx1* null mice (Vermot et al. 2003). Cardiac anomalies included PTA, abnormalities of the aortic arches including a right sided aortic arch and an aberrant right subclavian. In addition to the defects of the third to sixth branchial arch arches, the mice had hypoplasia of the thymus and parathyroid glands (Niederreither et al. 2001).

Given the complexities of retinoid signalling, it could be that exogenous retinoids affect *Tbx1* expression, and that *Tbx1* is an upstream regulator of certain factors involved in retinoid metabolism. Regulation may be direct via retinoic acid response elements (RAREs) or indirectly via other transcription factors.

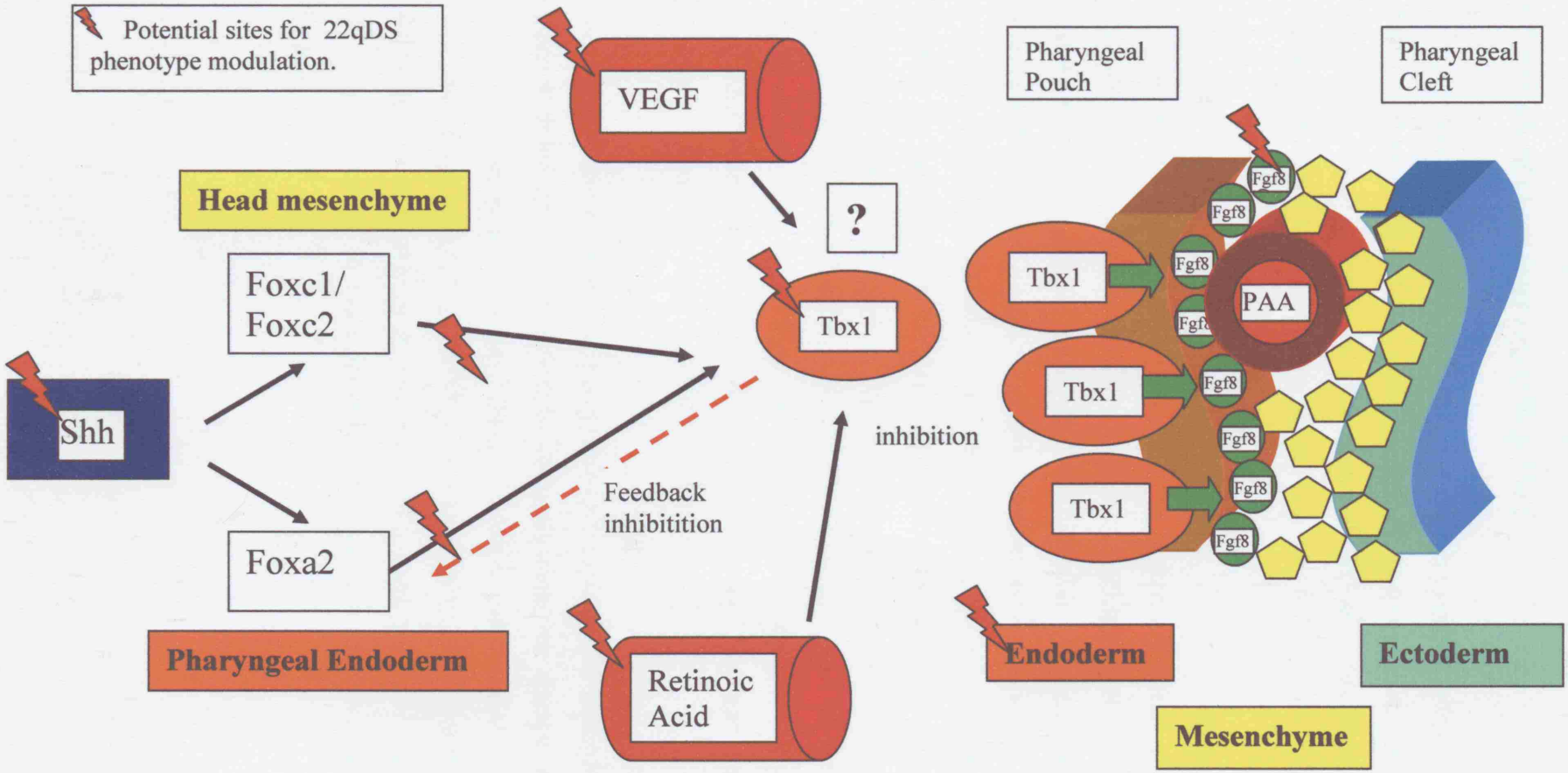


Figure 1.6.4
A proposed model for mechanisms of regulation of Tbx1.

In vitro studies have shown that *Tbx1* is regulated by *Foxa2* (with feedback inhibition), *Foxc1* and *Foxc2*. These genes are controlled by *Sonic Hedgehog (Shh)*. Tbx1 expression is also modified by *VEGF* and Retinoic Acid levels. *Fgf8* is a downstream target of *Tbx1* in the pharyngeal endoderm. PAA: pharyngeal arch artery. See text for references and details.

1.6.5 Unresolved Questions regarding Tbx1

As *Tbx1* is a transcription factor, it is likely that it has many protein binding partners that interact and affect either DNA binding activity or transcriptional activity. This has found to be the case for other T-box genes such as *Tbx5* and *Tbx4*, that have recently been shown to bind to *LMP-4* (Krause et al. 2004).

Although *TBX1* is presently the strongest candidate for many of the features of 22q11DS, animal models for 22q11DS suggest that haploinsufficiency of *TBX1* cannot solely account for the entire clinical presentation of 22q11DS. For instance the *Tbx1* deficient mouse does not show prepulse inhibition (Scambler-personal communication). It remains possible that the cumulative effect of several genes within the 22q11 region results in the full 22q11DS phenotype. These include *Hira* and *Ufd1l* which are expressed in the pharyngeal arches and are essential for murine survival, and *Crkl*, which causes a 22q11DS phenotype in homozygous-null mice. In order to unravel pathways involved in the aetiology of 22q11DS it is important to search for gene expression changes that result from the deletion. Two large remaining unresolved questions include establishment of the downstream targets of *Tbx1* and identification of the gene changes that result from the 22q11 deletion.

Although *TBX1* mutations in non-deleted patients with a 22q11DS phenotype have now been reported, it remains to be seen whether these mutations are disease causing or not. It also leaves unanswered questions about the many patients that phenocopy the syndrome, but do not have a 22q11 deletion or a mutation in *TBX1*. For these patients it is possible they have mutations in pharyngeal expressed genes, or they could have submicroscopic deletions in other regions of the genome involving genes in *TBX1* signalling pathways.

1.7 Aims and objectives of project and overview of thesis.

Generally this project aimed to identify potential genes within the *Tbx1* pathway. The specific aims were several fold. Firstly potential protein binding partners of the C and the N terminal regions of mouse Tbx1 were sought by screening a mouse embryonic day 9.5/10.5 library through the use of the yeast two hybrid system. Chapter 3 describes this work and a potential interactor Dynein light chain 1, cytoplasmic was identified.

Secondly, identification of potential downstream targets of *Tbx1* was possible using murine RNA studies. Breeding of the *Df1* mouse colony (Lindsay et al. 1999a), supplied by our Baylor collaborators E.Lindsay and A.Baldini, was used to investigate downstream targets of mouse *Df1* genes including *Tbx1*. This study was made possible by means of an in-house Affymetrix oligonucleotide microarray facility in collaboration with Dr Mike Hubank. Upregulated and downregulated genes were examined in greater detail by using a RTQPCR and whole mount *in situ* hybridisation approach to confirm or refute their role as targets. Identification of *Crkl* and *FoxA2* as potential interactors has since been confirmed by published work from other groups (Guris et al. 2001; Hu et al. 2004; Guris et al. 2006). This work is described in Chapters 4 and 5 and was reported in 2005 (Prescott et al. 2005a).

Thirdly in-house access to the *Splotch* mouse model (courtesy of Professor Copp) which has phenotypic overlap with 22q11DS, allowed an expression analysis of genes in a homozygous system. Similar techniques were employed to the *Df1* microarray study, except genes dysregulated at a greater magnitude necessitated a modified statistical approach. This work is described in Chapter 6. Known *Pax3* targets as well as potential novel interactors were identified.

Forthly, access to an archive of patient DNA samples that show some of the features of 22q11DS but did not have a FISH detectable 22q11.2 deletion was utilised through collaboration with Dr Nigel Carter at the Sanger centre to search for genomic changes in other regions that may harbour genes in the *Tbx1* pathway. Three chromosomal imbalances were identified, one of which was novel. Follow up work confirmed this imbalance (Prescott et al. 2005b). The results of this study are described in Chapter 7.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Reagents

AnalaR grade reagents obtained from either British Drug Houses (BDH) or Sigma Aldrich were used except where indicated otherwise. Glassware, solutions and media were autoclaved at 105kPa (15psi), 121°C for 20 minutes as required. Water was purified using a MilliRo 15 Water Purification System (Millipore SA), or further purified using a Milli-Q Reagent Grade Water Ultrafiltration System (Millipore), and sterilised by autoclaving as necessary.

Tris-buffered phenol was obtained from Fisons Scientific Equipment. Agarose and TEMED (N,N,N',N'-tetramethylethylenediamine) were obtained from Gibco-BRL Life technologies. Acrylamide solution 40% w/v, (29:1 acrylamide:N,N@-methylene bisacrylamide) and urea were from Bio-Rad Laboratories. Amberlite resin (50mg/ml) was from Perkin Elmer. Alconox detergent was from Alconox IncBacto-agar, bactotryptone, agar, nitrogen base without amino acids and yeast extract were from Difco laboratories. Absolute alcohol was from Hayman Ltd. Ultrapure dNTP's were supplied by Pharmacia Biotech or Amersham Biosciences. Subcloning efficiency DH5α Competent Cells were obtained from Life Technologies. Deionised formamide was from Eastman Kodak Co. Penicillin and streptomycin were from Gibco-BRL Life Technologies. Fetal bovine serum (FBS) was from Imperial Laboratories Ltd. Amberlite resin and blue dextran/EDTA (50mg/ml and 25mM respectively) were from Perkin Elmer Applied Biosystems Division. Proteinase K was obtained from QIAGen Ltd. NaOH pellets were from Fisons Scientific Equipment. Pellet Paint® NF Co-Precipitant was supplied by Novagen. 10x MegaBACE™ LPA Buffer was from Amersham Biosciences. TRIzol was from Invitrogen Life Technologies and glycogen was from Ambion.

For the Affymetrix microarray experiment: Bovine Serum Albumin (BSA) solution (50 mg/mL) was from Invitrogen Life Technologies, Herring Sperm DNA was from Promega Corporation, GeneChip Eukaryotic Hybridisation Control Kit (contains Control cRNA and Control Oligo B2) was from Affymetrix. 5M NaCl, RNase free,

DNase-free was from Ambion, MES hydrate SigmaUltra, MES Sodium Salt, DMSO and Goat IgG, Reagent Grade were from Sigma-Aldrich. Surfact-Amps 20 (Tween-20), 10% was from Pierce Chemical. Phycoerythrin Streptavidin was from Molecular Probes and 20X SSPE (3M NaCl, 0.2M NaH_2PO_4, 0.02M EDTA) was from BioWhittaker Molecular Applications. Phenol/chloroform/isoamyl alcohol and RNAlater™ were from Ambion.

For the *in situ* hybridisation experiments: DIG RNA labelling Mix, anti-Digoxigenin-AP, Fab fragments and NBT/BCIP ready-to-use tablets were from Roche. Tween-20, yeast RNA, heparin, gluteraldehyde were from Sigma. CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]-1propanesulfonate hydrate) was from BDH, PBS (phosphate buffered saline) was from Invitrogen, formamide was from Promega, and sheep serum was from Gibco.

2.1.2 Materials

In addition to standard laboratory items, use was made of the following special items. Aerosol resistant barrier tips were obtained from Molecular Bioproducts or CLP. Falcon 15ml tubes, 50ml tubes, 2059 polypropylene tubes and tissue culture flasks were from Becton Dickinson Labware. Culture plates were from Bibby Sterilin Ltd. Glassware was obtained from Pyrex® (USA) and Schott Duran (Germany). Plastic ware was sourced from Becton Dickson Lab ware or Bibby Sterilin Ltd. 0.2ml PCR tubes from ABgene®. Polycarbonate 96 well microplates OMNIPLATE 96 ThermoHybaid were from Hybaid Ltd or ABgene (Thermofast ® 96 skirted plates). Plate seals HB-TD-TAPE Micro sealing sheets were from Hybaid Ltd. PCR Miniflex 0.2mm flat tips were from Sorenson BioScience, Inc. Sterile loops, cryotubes and filter sterilisation units were acquired from Nunc Nalgene®.

DNA electrophoresed on agarose gels was visualised through ethidium bromide (Sigma-Aldrich) fluorescence detected at 300nm on a Chromato-Vue transilluminator (UVP, Inc.). Gels were photographed using a video cameraon and printed onto thermal paper using a thermal printer (Mitsubishi Electric Corporation). PCR amplification and cycle sequencing reactions were performed on a GeneAmp PCR System 2400 (Perkin Elmer Applied Biosystems) or Eppendorf Mastercycler gradient thermocycler (Eppendorf). DNA sequence analysis and microsatellite repeat analysis were performed

on an ABI PRISM™ 377 DNA Sequencer (Perkin Elmer Applied Biosystems Division) or a MegaBACE 1000 DNA Sequencer (Amersham Biosciences). Plastic and paper sharks tooth combs were supplied by Perkin Elmer Applied Biosystems Division. Phase lock gel was from Brinkman Instrument.

Assessment of RNA integrity was performed on an Agilent 2100 Bioanalyzer. Expression analysis was performed on Affymetrix Genechip® microarrays. Microarray target hybridisation was performed on the Affymetrix Fluidics station. RTQPCR was performed on the DNA Engine Opticon 2 System (M.J Research).

2.1.3 Commercial Kits.

Thermo Sequenase dye terminator cycle sequencing pre-mix kits were purchased from Amersham International. QIAquick PCR purification kits, QIAEX II Agarose Gel extraction kits, QIAprep Spin Miniprep and Maxiprep kits and RNeasy clean up kits were from QIAgen Ltd. RNA transcript labelling kit (synthesis of biotin labelled RNA by *in vitro* transcription) was purchased from Enzo Life Sciences. Nick translation kit for FISH probe generation was from Vysis.

2.1.4 Enzymes.

Restriction endonucleases were obtained from Gibco-BRL Life Technologies. Biopro™ DNA polymerase was from Bioline. Cloned Pfu DNA polymerase was from Stratagene. AmpliTaq Gold™ was obtained from Perkin Elmer Applied Biosystems Division. Biopro™ Taq polymerase was from Bioline. All enzymes were used with the appropriate 10x buffers according to the manufacturer's instructions. E. coli DNA Ligase, E. coli DNA Polymerase I, E. coli RNaseH, T4 DNA Polymerase, 5X Second-strand buffer and SuperScript™ II were all supplied by Invitrogen Life Technologies.

2.1.5 Nucleotide and protein size markers.

100bp and 1kb DNA ladders were purchased from Gibco-BRL Life technologies and were used for all DNA analysis. Wide range protein markers were purchased from Sigma and used for protein analysis.

2.1.6 Bacterial Strains.

DH5α: F$^-$, Φ80, *Lac*Z∇M15, ∇(*Lac*ZYA-*arg*F), U169, *deo*R, *rec*A1, *end*A1, *hsd*R 17(r_k^-, m_k^+), *sup*E44, λ-, *thi*-1, *gyr*A96, *rel*A1

2.1.7 Oligonucleotides

All oligonucleotide primers were synthesised to order by Genosys Biotechnologies Inc or QIAgen Ltd. Fluorescent dye labelled oligonucleotides were ordered from QIAgen Ltd.

Generation of cDNA from RNA was performed with random primer oligonucleotides (mostly hexamers) supplied by Invitrogen life technologies.

GeneChip T7-Oligo(dT) Promoter Primer Kit:
5′ - GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT)24 - 3′
50 μM, HPLC purified was from GENSET.

Primers used for general PCR, RTQPCR, sequencing of clones and genotyping during this project are listed in the following tables.

Table 2.1.7.1

Primers designed to generate yeast two hybrid Gal4 fusion constructs. Italics: *EcoRI* restriction sites, underlined *SalI* restriction sites.

CONSTRUCT	PRIMER SEQUENCE	AMPLIFIED REGION
pGBDU-C3 Nterm.mtbx1	F 5' cc*ggaattc*atctccgccgtgtctagtcc 3' R 5'accggtcgacctgcacgctcacgctggcc 3'	299bp
pGBDU-C3 Cterm.mtbx1	F 5' cc*ggaattc*aaccaccggcccggagcg 3' R 5' accggtcgacctatctggggcagtagtcgta 3'	642bp
pGBDU-C3 FLmtbx1	F 5' cc*ggaattc*atctccgccgtgtctagtcc 3' R 5' accggtcgacctatctggggcagtagtcgta 3'	1463bp

Table 2.1.7.2

General vector sequencing primers

VECTOR	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	ANNEALING TEMPERATURE (^{0}C)
pGBDU-C3	tgcctctaacattgagacag	cacagttgaagtgaacttgc	59
pVP16	gagtttgcgcagatgttt	gttgtaaaacgacggccagt	61

Table 2.1.7.3

Primers used for genotyping embryos

GENE	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	ANNEALING TEMPERATURE (^{0}C)
Pax3	cctcggtaagcttcgccctctg	cagcgcaggagcagaaccaccttc	60
Df1: 3 primer PCR	cagagttctgacttctgcacactaa		57
	tgggcaattgtttaatcttcc		
	tctttgtcagcagttcccttt		
Tbx1: 3 primer PCR	tcgactagagcttgcggaac		65
	agtctggggactctggaagg		
	aaggcagatcctgctacacc		

Table 2.1.7.4

Primers used for RTQPCR

GENE	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	ANNEALING TEMPERATURE (^{0}C)
Gapdh	ttcaccaccatggagaaggc	ggcatggactgtggtcatga	57
Tbx1	cgacaagctgaactgacca	caatcttaagctgcgtgatcc	57
ES2	cttccctgatgtggagaagcta	aggagtttcaaaggtggctgga	57
Cdc45	gttcaagcataaattcctggcc	ctagaccaaggtatagcttgtc	57
Comt	gtggacacattagacatggtc	ctcacatacgccaggaagtc	57
Nlvcf	gagagcagaacctctgcgga	cttcaataggaatcagttcttgg	59
T10	cgcccgtggtagaggtgaa	gctcaggtcagctgctatga	57
Txnrd	caacaagtcacagtaggctc	cttcatatgcacaggtgatgc	57
Ufd1L	agatcggatgacacactgcg	caggcgaagtctctcaatgc	57
Hira	gttgtgatggaatatgtctc	cttccacaccataatcagtttg	57

GENE	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	ANNEALING TEMPERATURE (^{0}C)
Mef2c	agcaagaacacgatgccatca	tactcagtaccatatgttgttga	57
Tcf20	tggaaatcgccagagagatga	caatctgcgtcaatggcacat	57
Ppir2	tgttgaagagtcaagtcaagga	ttgtatttgacaacgaccgca	57
IA-1	gtgcgtccggcctgctaga	ctccaccgaagcgaagcgaa	59
Crkl	gtgtctcgcactacatcatcaa	gctgagacagaacccactgg	59
Tbx2	cttcatcgctgtcactgccta	ccatcgctccggcttaca	59
FoxA2	ctaagcgagctaaagggagca	ggtggtgaaggcgtaatggt	59
Pkrir	cctagtgtggagctcttgcc	cagcaatcaagtgaggagcag	59
Rarg	gctgggcaacaaacaaagta	atccgcagcattaggatgtcca	61
Conn45	ggaaagcaacaaacaaagta	ctcagcaggcgagtcagga	59
Wnt5a	ccaggaggactccgcagcc	atgggcttcttcatggcgagg	61
Sox3	ctgtggtgtttgcggagaaac	gcaacctcactcagttctcga	59
DnaJb9	ctgcctcagagcgacaaatca	tccgactattggcatccaga	59
Gnbp1	ggtgccaccttccagctcta	ggcctttctagccagccagcagga	59
Shh	gaaagcagagaactccgtgg	gggacgtaagtccttcacca	57
p53	acctcactgcatggacgatct	gacactcggagggcttcactt	61
Met1	agctcctgcgcctgcaagaa	ggcacatttggagcagccca	61
Mad4	ccgaacaacaggtcttcacaca	acgcttcagaaggctcagagta	61
Sox11	gttcacgtattgagaggcgc	caatttctctgccacatct	57
Pax3	gatggcatcctgagtgagcga	gctaaaccagacctgcactcg	61
Aggrecan	acggcatcgaggacagcgaa	gatggcgctgttctgtaggc	59
Hmx1	gaggatccggagcaggcg	acctccgtagccgccgtg	57
Lunatic fringe	gagacctggatctcgcgcc	atcggtcatactccacagcca	59
Pias	ctctttgactggctgtcggt	acccacaacgtgatctgctca	57
Pak3	tcagactttgagcatacgattca	cttggagtcgtagaacttgag	59
Sox11	gttcacgtattgagaggcgc	caatttctctgccacatct	57
Gata5	ggagtgcgtcaactgtggag	gagcagcataggccggatct	59
Pax3	gatggcatcctgagtgagcga	gctaaaccagacctgcactcg	61
Aggrecan	acggcatcgaggacagcgaa	gatggcgctgttctgtaggc	59
Hmx1	gaggatccggagcaggcg	acctccgtagccgccgtg	57
Lunatic fringe	gagacctggatctcgcgcc	atcggtcatactccacagcca	59
Pias	ctctttgactggctgtcggt	acccacaacgtgatctgctca	57
Pak3	tcagactttgagcatacgattca	cttggagtcgtagaacttgag	59
Gcm2	cagtggtgactgactggcaa	cctaagaggctacttgatgac	57

Table 2.1.7.5

Primers used for genotyping (8q11.2). AT: annealing temperature.

MARKER	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	AT (^{0}C)	LENGTH	LABEL
D8S510	ctggcccaccagtctt	ggcaacgcatggacta	55.8	217-225	Tet
D8S1696	ctgccctagaaggctaactacc	tgccatttgctcactgagat	53	237-279	Fam
D8S1178	agctacttggnaggctga	aacattgggaatgggttatt	58.5	235-277	Tet
D8S1812	agaaggcaccactaatcc	tttggaaggcaaaaca	55	159-173	Tet
D8S1763	caatgaaagcctttatctacaggt	ttccattcacggagtaaaagt	55	157-167	Fam
D8S1113	atgaagatgaaccaggaa	ccctggactcatggtacttg	55	215-245	Tet
D8S512	tccagttttgctcccc	aagtgggatgagggcg	55	127-133	Fam
D8S544	gccattatgctgtcttgc	ctcttgtggaatttgaggg	55	132-140	Tet
D8S260	aggcttgccagataaggttg	gctgaaggctgttctatgga	55.8	187-213	Fam

Table 2.1.7.6

Primers used for genotyping (5q11). AT: annealing temperature.

MARKER	FORWARD PRIMER 5' → 3'	REVERSE PRIMER 5' → 3'	AT (^{0}C)	LENGTH	LABEL
D5S623	ttgcgtccccattagg	cccacactgtccatgc	55	102	Tet
D5S664	aattgttccagccactaccc	gccaccttctgagggg	58	199-143	Fam
D5S1969	agggaacctcacctgg	gacaagggctgggatg	62	217-269	Fam
D5S2068	acaatttagtggaggagcaaaa	atgtaagagtctttagaaacaggca	55	239-263	Tet
D5S1968	ggttctcaccacctcctg	atcattatgcttccccgac	55	188-222	Tet
D5S461	gcaacttaaggacatggcac	taagggtatcgggttggaag	55	180-192	Tet
D5S491	acacctggccccatct	ggctggacaatacacctcat	55	161-169	Tet
D5S2102	ccacaatggtgtcagcc	cctgcctctccagagaaac	55	116-124	Fam
D5S2076	catacacggtggactcattc	gctttcaacttaaaagatattgttg	55	103-127	Tet
D5S2037	tatgcgatgaaaggcaac	ggtaaggaactggactatctctg	56	278	Tet
D5S2035	tgaggttccattgagttacttc	cattacaaaactgtagtgctctgc	56	172	Fam
D21S264	gctcaatggctcaatttcc	ctgttggtaacacaaatgtattcc	56	250	Tet
D5S466	ctggttattggtcaacataccc	ataggtgatggcaactctgc	55	175-187	Tet
D5S407	tggtttagagaatttgcccc	ctgtcatttgttgttcattggaagt	55	107-127	Tet
D5S2507	ctttctaataacattgggggg	gccatagcctttcaactcaa	55	224-244	Fam

Table 2.1.7.7

Details of Western antibody

Antibody	Supplier	Host	Concentration
			Western Analysis
α-Gal4	Santa Cruz	Rabbit	1/500

2.1.8 Nucleotide Size Markers

The 50bp and 1kb DNA ladders were from Gibco-BRL Life Technologies. The MegaBACE™ ET400-R and MegaBACE™ ET550-R size standard were from Amersham Biosciences.

2.2 Solutions, Buffers and Media

2.2.1 DNA related solutions

2.2.1.1 Solutions for DNA preparation

Buffer P1	50mM Tris-HCl, pH 8.0; 10mM EDTA; 100μg/ml RNaseA
Buffer P2	200mM NaOH, 1% SDS
Buffer N3	3M KAc pH 5
PE	Concentrate (QIAgen)
Buffer EB	10mM Tris Cl, pH 8.5
Buffer P3	3M KAc pH 5
QBT	750mM NaCl, 50mM MOPS (pH 7.0), 15% isopropanol, 0.15% Triton-X-100
QC	1M NaCl, 50mM MOPS (pH 7.0), 15% isopropanol
QF	125mM NaCl, 50mM Tris-HCl (pH 8.5), 15% isopropanol

2.2.1.2 DNA Electrophoresis buffers

1 X TAE buffer	40mM Tris-Acetate, 1mM EDTA pH 7.0
1 X TBE buffer	89mM Tris-borate, 1mM EDTA pH 8.3

2.2.1.3 DNA Gel loading buffers

Bromophenol Blue	50μl 5% bromophenol blue, 240μl 80% glycerol, 710μl H_2O

2.2.1.4 Restriction Enzyme Buffers

Buffers used in restriction enzyme digest were enzyme specific and supplied at 10 X concentration by Gibco BRL. All were stored at -20°C.

1X BRL restriction buffer	50mM Tris-HCl (pH8.0), 10mM $MgCl_2$, 0-100mM NaCl, 0-100mM KCl

2.2.1.5 Taq PCR Buffer

Taq PCR buffer was supplied by Bioline at 10 X concentration and stored at -20°C.

10X *Taq* buffer	670mM Tris-HCl, pH 8.8, 160 mM $(NH_4)_2SO_4$, 1% Tween-20

2.2.1.6 Pfu PCR buffer

Pfu PCR buffer was supplied by Stratagene at 10 X concentration and stored at -20°C.

10X *Pfu* PCR buffer	200mM Tris-HCl pH 8.8, 20mM Mg SO_4, 100mM KCl, 100 mM $(NH_4)_2SO_4$, 1% Triton X-100, 1 mg/ml nuclease free BSA

2.2.2 RNA related solutions

2.2.2.1 RTQPCR

Superscript II reverse transcriptase 5x buffer	250mM Tris-HCl, pH 8.3 at room temperature, 375 mM KCl; 15 mM $MgCl_2$.

SYBR Green PCR Master Mix contains

HotStarTaq DNA Polymerase
QuantiTect SYBR Green PCR Buffer Containing:
Tris·Cl, KCl, (NH4)2SO4, 5 mM $MgCl_2$,
pH 8.7 (20°C)
dNTP mix: Contains dATP, dCTP, dGTP, and dTTP/dUTP;
Fluorescent dye: SYBR Green I
RNase-free water

2.2.2.2 Whole Mount In situ Hybridisation Buffers

10x transcription buffer	400mM Tris-HCL, pH 8.0 (20°C), 60mM $MgCl_2$, 100mM dithiotreitol (DTT), 20mM spermidin.
NBT/BCIP tablets	After the addition of double distilled water: 0.4mg/ml Nitro blue tetrazolium chloride (NBT); 0.19 mg/ml 5-Bromo-4-chloro-3-indolyl phosphate, toluidine salt (BCIP); 100mM Tris buffer, pH 9.5; 50mM $MgSO_4$.
PTw	PBS, 0.1% Tween-20
SSC	0.3 M sodium citrate, pH approx. 5.3, containing 3 M NaCl
10xTBST	4g NaCl; 0.1g KCl; 12.5 ml 1M Tris-HCl pH 7.5; 6.5g Tween-20; 32ml H_20.
NTMT	1ml 5M NaCl; 2.5ml 2M Tris HCl (pH 9.5); 1.25ml 2M $MgCl_2$; 5ml 10% Tween-20; 40.25ml H_20.

Whole mount *In situ* Hybridisation solution

Component (stock concentration)	**Final concentration**	**Volume to add**
Formamide	50%	25ml
salt sodium citrate (SSC) (20x)	1.3xSSC	3.25ml
EDTA(0.5M, pH 8.0)	5mM	0.5ml
Yeast RNA (5mg/ml)	50µg/ml	50µl
Tween-20 100%	0.02	100µl
CHAPS (10%)	0.005	2.5ml
Heparin (50mg/ml)	100µg/ml	100µl
H_2O		17.5ml
Total		**50ml**

2.2.2.3 Affymetrix Buffers

5X RNA Fragmentation Buffer made with RNase free H_2O	200 mM Tris-acetate, pH 8.1, 500 mM KOAc, 150 mM MgOAc
2X Hybridisation Buffer	Final 1X concentration is 100 mM Methanesulfonic acid (MES), 1 M [Na+], 20 mM EDTA, 0.01% Tween 20
Stringent Wash Buffer	100 mM MES, 0.1 M [Na+], 0.01% Tween 20
Non-Stringent Wash Buffer	6X SSPE, 0.01% Tween 20
2X Stain Buffer	Final 1X concentration: 100 mM MES, 1 M [Na+], 0.05% Tween 20

2.2.3 Protein analysis solutions

2.2.3.1 12% SDS Polyacrylamide gel mix components

Tris-HCl	1.5M, pH 8.8	2.5ml
Acrylamide Mix	30%	4ml
Sodium dodecyl sulphate (SDS)	10%	0.1ml
Ammonium Persulphate	10%	0.1ml
H_20		3.3ml
Temed		10µl

2.2.3.2 General Buffers

10 X SDS PAGE Running Buffer	0.25M Tris base, 2M glycine, 1% SDS

2.2.3.3 Western analysis buffers

10 X TBS	0.5M Tris-HCl (pH 8.0), 1.5M NaCl
TBS-T	1X TBS, 0.05% Tween-20
Blocking buffer	1X TBS; 5% non-fat milk; 0.05% Tween-20
Wash buffer	1X TBS; 0.5% Tween-20
Stripping Buffer	62.5mM Tris-HCl, pH 6.8; 2% SDS; 100mM β-mercaptoethanol

2.2.3.4 Ponceaus staining

PonceauS stain	0.3% PonceauS in 5% Trichloroacetic acid (Sigma)
PonceauS de-stain	H_20

2.2.4 Yeast Analysis

Yeast Strain

pJ69-4A MATa, *gal4Δ*, *gal80Δ*, *trp*1-901, *leu*2-3, *ura*3-52, *his*3-200 GAL2-ADE2, LYS2::GAL1-HIS3, met2::GAL7-LacZ.

2.2.4.1 Yeast media

YPD	20g/l Bacto-peptone, 10g/l yeast extract, 2% glucose
YPAD	YPD + 0.002% adenine
Synthetic Dextrose Minimal media	6.7g/l yeast nitrogen base (without amino acids), 2% glucose

Table 2.2.4.1

Amino acids used in Yeast Two Hybrid system

Amino acids according to selection criteria for Yeast 2 Hybrid system	Final concentration in medium (mg/litre)
Adenine Sulphate	20
L-Tryptophan	20
Uracil	30
L-Histidine HCl	20
L-Methionine	20
L-Leucine	60

Plates: Appropriate medium with addition of 20g/l agar

2.2.4.2 Yeast manipulation solutions

10X LiAc — 1M Lithium Acetate

10X TE — 100mM Tris-HCl (pH8.0), 10mM EDTA

1 X LiAc/1 X TE — 100mM LiAc, 10mM Tris-HCl (pH8.0), 1mM EDTA

1X LiAc/1X TE/ 40% PEG — 0.1M LiAc, 10mM Tris-HCl (pH8.0), 1mM EDTA, 40% polyethylene glycol (PEG)

Yeast lysis buffer — 10mM Tris-HCl (pH8.0), 2% Triton X-100, 1% SDS, 100mM NaCl, 1mM EDTA

Z-buffer — 60mM $Na_2HPO_4.7H_20$, 40mM $NaHPO_4.H_2O$, 10mM KCl, 1mM $MgSO4.7H_2O$, adjusted to pH 7.0

Cracking buffer — 125mM Tris-Cl, pH 6.8, 8M urea, 4% SDS, 10% glycerol, 0.05% bromophenol blue and 50µl/ml β-mercaptoethanol added fresh.

2.2.5 Bacterial Analysis

2.2.5.1 Bacterial Media

L-Broth (LB)	10g/l Bacto-tryptone, 5g/l yeast extract, 5g/l NaCl
Plates	L-Broth with the addition of 20g of agar

2.2.5.2 Antibiotics

Ampicillin	50mg/ml, filter sterilised stored at -20°C as 1000X stocks
Kanamycin	25μg/ml final concentration
Chloramphenicol	20μg/ml final concentration

2.3 Methods

2.3.1 Isolation and Purification of DNA

2.3.1.1 QIAgen QIAprep® Miniprep

1ml of a total 5ml overnight culture was harvested by centrifugation at 13000 rpm for 1 minute. Bacterial cell pellets were resuspended in 250μl of buffer P1, with subsequent of addition of 250μl of P2.These were then inverted 4-6 times to mix. 350μl of N3 was then added and samples inverted 4-6 times to mix. Precipitates were harvested by centrifugation at 13000rpm for 10 minutes. Supernatants were bound to QIAgen spin miniprep columns by centrifugation at 13,000 rpm for 1 minute and washed with 750μl of PE using centrifugation as above. The columns were centrifuged to remove excess PE and DNA eluted from the column by incubation with 50μl of EB for 1 minute, followed by centrifugation at 13,000 rpm for 1 minute.

2.3.1.2 QIAgen QIAprep® Maxiprep

Large scale DNA purification was conducted according to manufacturer's protocol, QIAGEN® purification handbook. 200ml of bacterial culture was harvested by centrifugation at 6000 rpm for 15 minutes at 4°C. Bacterial pellets were resuspended in 10ml of buffer P1 containing RNase A. 10ml of buffer P2 was added and the mixture inverted 4-6 times for mixing and incubated at room temperature for 5 minutes. 10ml of chilled buffer P3 was then added; the sample was mixed by inverting 4-6 times, and incubated on ice for 20 minutes. Cell debris was harvested by centrifugation at 10,000rpm for 30 minutes at 4°C and the supernatant re-centrifuged as above for 15 minutes. During this period, QIAgen-tip 500 columns were equilibrated by applying 10mls of buffer QBT, and allowing the columns to empty by gravity flow. The cleared lysates were added to the columns and allowed to enter the resin by gravity flow. The columns were then washed twice with 30mls of buffer QC and the DNA eluted from the columns using 15mls of buffer QF. The DNA was precipitated by the addition of 10.5mls of room temperature 100% isopropanol, and pelleted by centrifugation at 10,000rpm for 30 minutes at 4°C. The supernatant was carefully decanted and the pellet washed with 5mls of 70% ethanol and centrifuged at 10,000rpm for 30 minutes at 4°C. DNA pellets were air dried and resuspended in 200µl of 10mM Tris-Cl, pH8.5.

2.3.2 Analysis of DNA and RNA

2.3.2.1 Quantification of Nucleic Acid Concentration

Nucleic acid concentration was quantified by measuring the absorbance of a solution at A_{260} and using the following equations,

$$\text{DNA concentration } \mu g/\mu l = OD_{260} \times \frac{\text{dilution} \times 50}{1000}$$

$$\text{RNA concentration } \mu g/\mu l = OD_{260} \times \frac{\text{dilution} \times 40}{1000}$$

The A_{260}:A_{280} ratio was determined to asses the amount of protein contamination.

2.3.2.2 General PCR amplification of DNA

PCR reactions were conducted in 25 to 100µl volumes containing 1X polymerase buffer, a 10mM dNTP mix, 50mM $MgCl_2$ and 25µM of each forward and reverse primers. Between 30 and 90ng of template DNA and 0.2-1 unit of DNA polymerase were added to the mix. PCR products subsequently used for subcloning and protein expression studies were amplified using the proof reading polymerase *Pfu*, whist for other applications the non proof reading Bioline *Taq* polymerase was used. The following PCR conditions were used for both *Pfu* and Bioline *Taq* polymerase,

1 Cycle:	Denaturation	94 °C	2-5 minutes
35 Cycles:	Denaturation	94 °C	1 minute
	Annealing X °C		1 minutes,
	where X= appropriate annealing temperature		
	Elongation	72 °C	1-3 minutes

2.3.2.3 PCR amplification of DNA for microsatellite analysis

PCR amplification of microsatellite repeats was carried out in accordance with standard protocols. AmpliTaq Gold (Perkin Elmer Applied Biosystems) was used. The hot start has got the advantage of improving amplification by lowering non-specific background and non-specific primer annealing, as the enzyme is inactive at room temperature. Amplification reactions contained 1x GeneAmp PCR Buffer II, 2.5mM $MgCl_2$, 0.25mM of each dNTP, 1µM of each primer, 0.3U AmpliTaq Gold and approximately 100ng template DNA in a final volume of 10µl. (10x GeneAmp PCR buffer II contains 100mM Tris-HCl (pH 8.3) and 500mM KCl). Initial incubation (to activate the enzyme) and denaturation was at 95°C for 12 minutes, followed by 10 cycles of denaturation at 94°C for 15 seconds, annealing at 53-55.5°C for 15 seconds and extension at 72°C for 30 seconds, and a further 20 cycles of denaturation at 89°C for 15 seconds, annealing at 53-55.5°C for 15 seconds and extension at 72°C for 30 seconds.

To assess the success with which the repeats had been amplified, a 5µl aliquot of each completed PCR reaction was first electrophoresed on a 2% agarose gel. Depending on

the quantity of the product obtained, the remaining sample was diluted 1 in 5 or 1 in 10 in sterile purified water.

2.3.2.3 Primer Design for genotyping

The primer sequences were taken directly from public databases (uniSTS at http://www.ncbi.nlm.nih.gov/genome/sts). 6-FAM and TET 5' labelling (blue and green respectively) was preferentially used.

2.3.2.4 Primer design for direct sequencing

For sequencing of exons, oligonucleotides primers were designed from genomic sequence with the aid of the program PRIMER3. Since these primers were used for direct sequencing of the PCR product after the initial PCR amplification, they were designed to be far enough 5' and 3' from the region of interest so that good quality sequence of the entire exon and splice site could be obtained.

2.3.2.5 Agarose Gel Electrophoresis

DNA fragments in the size range of 100bp to 10Kb were resolved by horizontal agarose gel electrophoresis. Between 1 and 2% agarose gels containing 0.5μg/ml ethidium bromide were used depending on the size of the DNA fragment to be resolved (0.5-1g/100mls agarose dissolved in 1 X TAE buffer). Gels were run in 1 X TAE buffer, with time and voltage being dependent on the size and resolution required of the DNA fragment. Agarose gels were visualised by exposure to UV light.

2.3.2.6 Purification of PCR Products

PCR products were purified using the Exonuclease I, Shrimp Alkaline Phosphatase PCR purification protocol (Amersham), according to the manufacturer's instructions. This protocol utilizes two hydrolytic enzymes, Exonuclease I and Shrimp Alkaline Phosphatase to remove unwanted dNTPs and primers. The Exonuclease I degrades residual single-stranded primers and any extraneous single-stranded DNA produced by PCR. The Shrimp Alkaline Phosphatase hydrolyses remaining dNTPs from the PCR mixture, which would interfere with the sequencing reaction. The Exonuclease I/ shrimp alkaline phosphatase mixture is added directly to the PCR product, incubated at 37°C

for 15 minutes for treatment and then at 80°C for further 15 minutes to inactivate the enzymes.

A second method of PCR purification was gel extraction. PCR products were first resolved on a 1% low melting point agarose gel and then the band excised. The excised DNA was purified using QIAquick Gel extraction kit (QIAgen Ltd) according to manufacturer's protocol.

2.3.2.7 Restriction Enzyme Digest of DNA

DNA was digested according to manufacturer's guidelines using the appropriate reaction buffers for a given enzyme. Digestion reactions were allowed to proceed for 2 to 3 hours at the appropriate reaction temperature.

2.3.2.8 Subcloning of GAL4 fusion products for yeast two hybrid analysis

Three constructs were designed to represent the N terminus (nucleotides 4-333), C terminus (825-1467), and full length (4-1467) mouse Tbx1. These coding sequences were cloned in frame with a cassette encoding the Gal4 DNA binding domain in order to generate mTbx1-Gal4 fusion proteins.

Primers containing *EcoRI* and *SalI* sites were designed to amplify the Nterm, Cterm and full length mTbx1 cDNA using Pfx (proof reading enzyme) and 20 cycles of PCR amplification. DNA was extracted form the PCR product using phenol extraction. The products were cloned into pGBDU (C1) (accession number U70021: Appendix 1). The PCR products and pGBDU were *SalI* digested and gel purified using a gel extraction kit (QIAgen). The products of gel extraction were digested with *EcoRI* and gel extracted. The N, C and full length mTbx1 products were ligated into pGBDU and transformed into *Escherichia coli* DH5α competent cells and plated onto LB agar containing ampicillin. 2 positive clones from each construct were recovered and grown in LB broth with ampicillin overnight at 37°C and plasmid DNA was isolated using a miniprep kit (QIAgen). The constructs were sequenced to ensure they were in frame with the yeast Gal4 activation domain, and that no mutations had been introduced in the PCR amplification process.

2.3.2.8.1 Ligation of DNA Insert into Vectors

Ligation reactions were conducted in a 10µl volume containing 1X ligation buffer and 1 unit of ligase. Vector and DNA insert were generally used at a 1:4 ratio respectively, but this varied depending on the size of vector and insert. Samples were incubated at 4°C overnight. The ratio of vector to insert was determined using the following equation,

ng of insert = (ng of vector x Kb of vector/ ng of vector) x moles insert

2.3.2.8.2 Transformation of Bacteria with Plasmid DNA

DH5α competent cells were removed from -70°C freezer and thawed on wet ice. 5µl of ligation reaction was gently mixed with 50µl of chemically competent bacteria cells and incubated on ice for 30 minutes. Cells were heat shocked at 37°C for 20 seconds followed by further incubation on ice for 2 minutes. Cells were subsequently rescued with 1ml of LB with further incubation at 37 °C for 2 minutes. 0.95ml of room temperature LB was added. Cells were shaken at 225 rpm for 1 hour at 37°C. Transformed cells were plated on LB agar plates with appropriate antibiotic resistance and incubated at 37°C overnight.

2.3.2.9 Automated DNA Sequencing

Sequencing was performed using BigDye terminator v3.1 (Applied Biosystems). Cycle sequencing was performed in a total volume of 10µl, using 0.1-0.3µg plasmid DNA or 100ng PCR product, 5pmol primer, 1µl sequenase and 2µl buffer. Cycle Sequencing reaction conditions: 96°C-20secs, 55°C-5 secs, 60°C-4min, 32 cycles. PCR products were cleaned with Sephadex. The 10µl samples were then denatured at 95°C for 2 minutes and injected into the MegaBACE sequencer (Amersham Biosciences) following the manufacturer's instructions. 3 matrix tubes were taken out from storage (4°C) at least an hour before use and centrifuged at 4000rpm for 4 minutes. A 96 well buffer plate was prepared by adding 200µl 1x linear polyacrylamide (LPA) buffer (Amersham Biosciences) to each well. Prior to prerunning, capillaries were first rinsed with sterile purified water and then filled with matrix buffer. After rinsing the capillaries again with sterile water, the 96-sample plate containing the sequencing products was loaded into the cassette and the samples were injected at 3000V for 40 seconds. The

sample plate was then replaced with the buffer plate and the run was started at 10 000V for 70-100 minutes. DNA sequences were analysed using BioEdit software.

2.3.2.10 Genotyping using MegaBACE™ 1000 DNA Sequencer

Analysis of fluorescent microsatellite repeats was performed on the MegaBACE™ 1000 DNA Sequencer (Amersham Biosciences), a high-throughput fluorescence-based DNA analysis system utilizing capillary electrophoresis with 96 capillaries operating simultaneously. PCR products were adequately diluted and 2µl of the diluted product was added to the loading mix containing 0.5µl MegaBACE™ ET400-R (or MegaBACE™ ET550-R depending on size of products) size standard and 7.5µl sterile purified water into a polycarbonate 96 well sample plate. MegaBACE™ ET size standards consists of 20 (ET400-R) and 22 (ET550-R) double stranded DNA fragments in which one strand is end-labelled with an energy transfer dye. They allow precision sizing of DNA fragments up to 400 or 550 bp. The 10µl samples were then denatured at 95°C for 2 minutes and injected into the MegaBACE sequencer following the manufacturer's instructions. Results were analysed with the aid of MegaBACE™ Genetic Profiler software (Amersham Biosciences) which allows automation of the analysis of di- and tetranucleotide repeats in a Windows™ NT format.

2.3.3 Yeast Two Hybrid Analysis

2.3.3.1 Genotype of the Yeast Strain pJ69

Yeast strain PJ69-4a was used for GAL4 based yeast two hybrid experiments. Genotyping of PJ69-4a was performed prior to yeast two hybrid analysis to ensure that the yeast strain had not undergone spontaneous mutation. PJ69-4a was initially propagated on complete medium (YPAD) plates, followed by sequential streaking onto appropriate selective SD minimal media plates with sequential omission of His, Trp, Ade, Ura, Met and Leu to confirm that the yeast genotype was correct. Plates were incubated at 30°C for 3 days and once the correct genotype was confirmed, a 20% yeast glycerol stock was made, and stored at -70°C.

2.3.3.2 Preparation of Competent Yeast Cells

The PJ69-4a host yeast strain (James et al. 1996) was plated onto YPAD-agar plates (see Table 2.2.4.1 for media concentrations) and grown for 48-72 hours at 30°C until visible colonies were apparent. One colony was then picked and grown overnight in 5ml YPAD medium in a shaking incubator at 30°C. Growth density was assessed by measuring the optical density (OD) at 600nm using a spectrophotometer. 1-3ml was used to inoculate 60ml YPAD culture to an OD of 0.1 and left at 30°C with shaking until the OD reached 0.5-0.7, when the yeast culture was considered to be in exponential growth.

2.3.3.3 Transformation of Yeast Cells

Once the OD of the yeast cells had reached 0.5-0.7 the cells were pelleted and resuspended in 20ml of sterile distilled water and re-spun to ensure no traces of YPAD were left. The cleaned pellet was suspended in 300µl 1X LiAc/TE.

Carrier DNA (10µg/µl of herring testes DNA; Clontech) was denatured by boiling at 99°C for 5mins. Then 100ng of each of the 3 constructs was added to 100µl PJ69-4a in 1XLiAc/TE with 50µg of the denatured carrier DNA and mixed by inversion. Each solution was mixed thoroughly and 1xTE/LiAc/40% polyethylene glycol was added to each tube and the mixtures were incubated at 30°C with shaking for 30 minutes.

To carry out the transformations, 70µl of DMSO (Sigma) was added to each tube and the PJ69-4a cells were heat shocked at 45°C for 15 minutes. The PJ69-4a cells were pelleted and washed with distilled water to ensure no traces of DMSO were left. The transformed cells were then plated onto SD-uracil and placed at 30°C for 48-72 hours for selection to take place. Six colonies were then picked and restreaked onto SD-uracil and allowed to grow for 48hours at 30°C.

2.3.3.4 Autoactivation Analysis

The autoactivation potential of each bait protein was tested prior to use in the yeast 2 hybrid system. This was to determine whether the bait alone was able to result in reporter activation. Six colonies were transformed with bait plasmid (Ura selection) and

streaked onto selective media plates composing of –Ura, -His + 2mM 3AT to test the histidine reporter and -Ura and -Ade to test the adenine reporter. Plates were incubated at 30°C for 2-3 days prior to analysis.

For histidine selection 3–aminotriazole (3-AT) at 2mM was added to the plates as the bait has some histidine activity on its own. 3AT acts to increase the amount of *His3* activity required for yeast survival.

2.3.3.5 Testing for protein expression of constructs

2.3.3.5.1 Protein Isolation

The transformed yeast was grown to an OD of 0.8-1.0 and 1.5mls of the culture was taken and spun in the centrifuge for 30seconds. The pellet was resuspended in 75µl of cracking buffer (see table 2.2.4.2) and then vortexed. The sample was then frozen in liquid nitrogen, thawed to 70°C and vortexed for 15 seconds. This was repeated twice.

2.3.3.5.2 SDS-Polyacrylamide Gel Electrophoresis

Glass plates were cleaned with detergent and wiped with 70% ethanol. The plates were sandwiched together separated by 1 mm spacers and clamped along the edge. A 12% SDS polyacrylamide gel was made (section 2.2.2.1) as expected protein product sizes were 28kD (Tbx1 N terminus) and 36kDaltons (Tbx1 C terminus). ddH_20-saturated isopropanol was layered on top of the poured gel while the polymerisation took place, to prevent bubbles. Once set, the isopropanol layer was removed, the top of the gel rinsed with ddH_20 and the stacking layer of the gel poured.

Samples were prepared for electrophoresis by adding an equal volume of 2× gel loading buffer, and denaturing for 10 minutes at 95 °C before loading onto the gel. Gels were run in 1× SDS-PAGE running buffer (section 2.2.2.2) at 120 V for approximately 2 hours, until required resolution was achieved. Molecular weight rainbow markers were used (Sigma).

2.3.3.5.4 Western Blotting

SDS-PAGE gels were blotted using a Mini-Protean II (Bio-Rad) blotting system. Gels were pre-equilibrated for 15 minutes in transfer buffer. Hybond-C nitrocellulose membrane (Amersham) was pre-wetted in ddH_2O followed by pre-equilibration in transfer buffer along with four pieces of 3MM paper. Gels were placed on two pieces of 3MM paper and then overlaid by nitrocellulose membrane. A further two pieces of 3MM were overlaid on the membrane. The entire assembly was sandwiched between two sponges and placed in a plastic transfer cassette. Transfer was conducted at 4°C in a tank containing transfer buffer at 30-50mA overnight. Following transfer, the apparatus was disassembled and the membrane stored in X 1 TBS at 4°C until immunodetection. Proteins were visualised on the membrane by staining with 0.2% Ponceau S solution for 2-3 minutes followed by de-staining with 0.3% Trichloroacetic acid.

2.3.3.5.5 Western Analysis

Membranes were washed with X 1 TBS prior to immunodetection. Membranes were blocked in blocking buffer for 1 hour at room temperature, or overnight at 4°C with gentle rocking. Following blocking, membranes were incubated with primary antibody (1/500 anti-Gal4 rabbit polyclonal IgG) for 1-2 hours at room temperature with gentle rocking. Membranes were washed for 1 hour with wash buffer at room temperature, which was changed approximately five times, and then incubated with horseradish peroxidase-conjugated secondary antibody diluted in blocking buffer at 1:5000 for 1 hour, at room temperature. Membranes were washed again with wash buffer as described above, followed by two 10 minute washes with 1XTBS. Equal amounts (solution 1 and 2) of ECL detection reagent, were mixed and applied to the membranes for 1 minute. Excess detection reagent was drained, membranes were covered in cling film and exposed to Xograph blue X-ray film for between 30 minutes to two hours.

2.3.3.6 VP16 library screening Library Screening

One yeast colony streak transfected with the GAL4-Tbx1-Nterminus fusion construct was inoculated into 20ml of SD-uracil medium and grown for 24hours at 30°C. The GAL4-Tbx1-Cterminus fusion construct was transformed using the low scale protocol where the same procedure was used as below but the volumes were divided by 10.

OD_{600} was measured and an aliquot was inoculated into 1000ml of SD-uracil to produce an OD_{600} of 0.1-0.2. After 5 hours the yeast cells were pelleted and the pellets were washed with TE several times by resuspending the pellet and centrifugation. The pelleted cells were resuspended in 20ml of 1xTE/LiAc. 1ml of 10mg/ml denatured herring tested carrier DNA was added and 250ug of 50:50 mouse E9.5/E10.5 VP16 library DNA. 140ml of 1xTE/LiAc/PEG was added.

The mixture was incubated for 30 minutes at 30°C with slow shaking. Cells were then heat shocked at 45°C with 17.6mls DMSO for 6 minutes, and rinsed to remove the DMSO by pelleting the yeast cells, resuspending in distilled water and re-centrifuging. After transformation, pellets were resuspended in 2000mls of pre-warmed YPAD media and left to recover for one hour. Cells were then pelleted and washed and resuspended in SD-uracil–leucine (uracil is the selectable marker for the pGBDU vector and leucine is the selectable marker for the VP16 cDNA library: Appendix 1). The cells were placed at 30°C with vigorous shaking for 9 hours overnight. Cells were then pelleted by centrifugation and suspended in 20 aliquots of 0.5ml each. Each aliquot was then plated onto SD-uracil-leucine-histidine+2mM 3AT plates. In addition serial dilutions were made and plated onto SD-uracil-leucine to calculate the transformation efficiency.

After 72 hours, interacting clones were patched onto SD-uracil-leucine-adenine plus 2mg/ml X-α-Gal (BD biosciences) plates. Only colonies that grew on both selection plates and that were blue in the presence of X-α-Gal were considered as possible interactors.

2.3.3.7 Recovery of Plasmids from Yeast

A 5ml culture of minimal media was supplemented with amino acids required for growth of pJ69 with omission of leucine as selection for the library plasmid VP16. This was inoculated with a single yeast colony and incubated at 30°C for 48 hours shaking at 225 rpm. Cultures were centrifuged at 2500rpm for 5 minutes. After decanting the supernatant and adding 50µl of lysis buffer cultures were treated overnight with 200 units (30units/µl) lyticase (Sigma) in order to digest the yeast cell walls. After 24 hours

the culture was spun down and the pellet was resuspended and purified using a QIAgen miniprep kit and eluted in 40µl of elution buffer.

2.3.3.8 Analysis of Positive Yeast Clones

5µl of yeast miniprep DNA was used as template in a PCR reaction using vector specific primers described in table 2.1.8.2. PCR samples were resolved on a 0.8% agarose gel, and if necessary the DNA purified using the QIAquick PCR Purification Kit (QIAGEN Ltd). Between 4-5µl of the PCR product was used as template for sequencing reactions as described in section 2.3.2.7. To verify the identity of each interacting clone, the resulting sequences was analysed for homology to other cDNAs using the BLAST search on the NCBI mouse genome database (http://www.ncbi.nlm.nih.gov/blast). Only in-frame inserts that did not appear on the list of known yeast 2 hybrid false positives (http://www.fccc.edu/research/labs/gelomis) were then considered as possible interactors.

2.3.4 Affymetrix Expression Microarray Analysis of Df1 and Pax3 E10.5 Embryos

2.3.4.1 Mouse strain and embryos

Df1 and *Tbx1* mice were obtained from Antonio Baldini, Houston, USA and bred in the Western laboratories, Institute of Child Health. The mice contain the 1.2Mb deletion on mouse chromosome 16B, syntenic to the human DiGeorge syndrome locus on 22q11.2 (Lindsay et al. 1999a). They were both bred on a C57Bl/6 background.

The Sp^{2H} mouse colony was in house and the study was done in collaboration with Dr D. Henderson from Prof. A. Copp's laboratory at the Institute of Child Health. The Sp^{2H} mutation involves a 32 base pair deletion in the homeodomain of the *Pax3* gene. This deletion leads to the formation of a stop codon leading to a truncated protein lacking a functional homeodomain and the carboxy region. They were bred on a C57/Bl6 background.

Adult mice were paired overnight and females were checked for the presence of a copulation plug in the morning. At midday, females with a copulation plug were considered to be at day 0.5 post coitum or embryonic day 0.5 (e0.5). Pregnant females of E10.5 were culled by cervical dislocation and the uterus was placed in Dulbecco's Modified Eagle's Medium (DMEM) for further dissection.

2.3.4.2 General Considerations for RNA handling

RNA samples were always stored at -80°C and kept on ice while in use. Care was taken to try and maintain and RNase free environment. Latex gloves were always used and changed frequently and wherever possible RNA work was done on a separate bench to DNA work. Sterile, disposable, polypropylene tubes were used throughout the procedure. Non-disposable plasticware was cleaned with RNAse-free water. Water was either treated with 0.1% DEPC (diethyl pyrocarbonate) or taken directly from the Milli-Q Reagent Grade Water Ultrafiltration System.

2.3.4.3 RNA isolation from E10.5 branchial arches.

Embryos were dissected from the decidua and extra embryonic membranes and rinsed in phosphate buffered saline (PBS). Embryos were staged (the somite number was noted) and the branchial arches were dissected as shown in the figure 2.3.4.3 and included the 1st,2nd, 3rd 4th and 6th branchial arches. Embryos were left at -20 in 0.8ml of RNA*later* until genotyped.

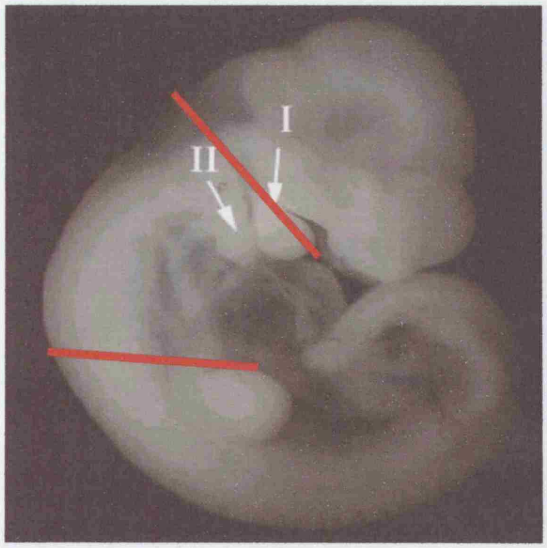


Figure 2.3.4.3

Wild type E10.5 mouse embryo showing branchial arch region dissected for microarrays and RTQPCR experiments. I and II depict 1st and 2nd branchial arches.

2.3.4.4 Genotyping of Df1 embryos

The yolk sac and remainder of each embryo was put in 50ul of PBS with 2µl proteinase K (10µg/µl) and left overnight at 55°C. The following day they proteinase K was inactivated by heating to 99°C for 10mins followed by a 12k spin for 15mins. The embryos were genotyped using the PCR conditions below. Df1 primers are shown in table 2.1.8.3.

Df1 genotyping PCR conditions

Template DNA	2µl	94°C-4mins
10x Taq Buffer	2.5µl	94°C-1min 57°C-1min 72°C-1min } x33
10mM dNTP	0.625µl	72°C-1min
50mM MgCl2	0.75µl	
3 Primer Mix (20pmol each)	1.0µl	
Taq polymerase	0.5µl	
ddH_20	17.625µl	

2.3.4.5 Genotyping of Sp^{2H} embryos

Methodology of DNA extraction from the yolk sac is the same as for Df1. Primers are detailed in table 2.1.8.3.

Template DNA	2µl	94°C-4mins
10x Taq Buffer	2.5µl	94°C-1min 60°C-1min 72°C-1min } x35
10mM dNTP	0.625µl	72°C-10min
50mM $MgCl_2$	0.75µl	
Primers		
Pax3b (50µM)	0.5 µl	
Pax3c (50µM)	0.5 µl	
Taq polymerase	0.5µl	
ddH_20	17.625µl	

2.3.4.6 RNA clean up and quantification

Once genotyped three embryos of the same genotype and stage were combined and were homogenised using a hand held eppendorf homogeniser. Total RNA was extracted using TRIzol (Invitrogen) and cleaned using the RNeasy Total RNA clean up kit

(QIAgen). RNA was then checked for quality using the Agilent bioanalyser. Quantification was done using spectrophotometry using the convention that 1 absorbance unit at 260nm equals 40µg RNA per ml. The absorbance was checked at 260nm and 280nm for determination of sample concentration and purity. Samples of low yield or quality were rejected for analysis.

2.3.4.7 Affymetrix Genechip protocol for synthesis of double-stranded cDNA from total RNA

2.3.4.7.1 Reverse transcriptase volumes for first strand cDNA synthesis reaction

Generally 8-16µg of total RNA was used for cDNA synthesis, but quantities of Superscript II reverse transcriptase were adjusted according to the following table.

Total RNA (µg)	SuperScript II RT 200U/µl (Gibco)
5-8	1.0
8.1-16	2.0
16.1-20	3.0

2.3.4.7.2 First Strand cDNA synthesis Components (Invitrogen)

1µl (100pmol) of T7-(dT)24 primer (100pmol/µl) and 5-20µg total RNA (final volume 20µl) were incubated at 70°C for 10 mins for primer hybridisation. 4µl 5X First Strand cDNA buffer, 2µl 0.1M DTT (10mM DTT) and 1µl 10mMdNTP (500µM) were added to the above tube and well and incubated at 42°C for 2 minutes to allow for temperature adjustment. Then SuperScript II reverse transcriptase is added at the appropriate amount (see above) and the samples were incubated at 42°C for 1 hour (final volume 20µl).

2.3.4.7.3 Second Strand cDNA synthesis Components

The first strand reaction tubes were placed on ice. The following reagents were then added: 91µl RNAse free H_20, 30µl 5X Second Strand Reaction Buffer, 3µl 10mM dNTP mix (200µM each), 1µl 10U/µl E.Coli DNA ligase (10U), 4µl 10U/µl E.Coli DNA polymerase I (40U), 1µl 2U/µl E. Coli RNase H (2U), final volume 150 µl.

The tube was gently tapped to mix and then incubated at 16°C for 2 hours in a cooling waterbath. 2µl (10U) T4 DNA polymerase was added and returned to 16°C for 5 minutes. 10µl 0.5 M EDTA was added to stop the reaction. The GeneChip Sample Cleanup Module (QIAgen) was used for sample clean-up as recommended by Affymetrix. Samples were eluted with 12µl Elution buffer.

2.3.4.8 Synthesis of Biotin labelled cRNA (In vitro Transcription: IVT)

The resulting cDNA was amplified using T7 RNA polymerase in the presence of biotin-UTP and biotin-CTP, so each cDNA will yield 50-100 copies of biotin labelled cRNA. The amount of cDNA used in the IVT reaction varied depending on the amount of starting total RNA (see table below).

Total RNA (µg)	Volume of cDNA to use in IVT (assuming 12µl resuspension volume for DNA)
5-8	10µl
8.1-16	5µl
16.1-20	3.3µl

The Enzo BioArray HighYield RNA transcript labelling kit was used as recommended by Affymetrix. Reagents were added in the order given in the following table:

RNA transcript *In vitro* Transcription (IVT) Labelling Reaction

Reagent	Volume
Template DNA	Variable to give 1µg
RNase free H_2O	Variable
10x HY reaction Buffer	4µl
10x Biotin-Labelled Ribonucleotides	4µl
10x DTT	4µl
10x RNase Inhibitor Mix	4µl
20xT7 RNA Polymerase	2µl
Total Volume	40µl

The samples were mixed carefully and placed in a 37°C water bath and incubated for 4-5 hours, with gentle mixing every 30-45 minutes. The samples were cleaned up using the columns supplied by the GeneChip clean up protocol. Samples were eluted in 2x10µl of RNAse free water.

2.3.4.9 Quantification of cRNA

Sample quality and quantity was established using the Agilent bioanalyser and spectrophotometric analysis. Samples were also run on a 1% agarose gel. An adjusted cRNA yield was calculated to reflect carryover of unlabeled total RNA. Using an estimate of 100% carryover, the formula below was used to determine adjusted cRNA yield:

Adjusted cRNA yield=RNA(m) – (total RNAi)(y)

RNA(m)= amount of cRNA measured after IVT (µg)

Total RNAi=starting amount of total RNA (µg)

y=fraction of cDNA reaction used in IVT

2.3.4.10 Fragmenting the cRNA for target Preparation

To minimise the amount of magnesium in the final hybridisation cocktail the cRNA was kept to a minimum concentration of 0.6µg/µl. 2µl of 5x fragmentation buffer was used for every 8µl of RNA plus water. The fragmentation buffer breaks down full-length cRNA to 35-200 base fragments by metal induced hydrolysis. The final concentration of RNA in the fragmentation mix was 0.5µg/µl to 2µg/µl.

The fragmentation buffer contained 40 mM Tris-acetate (pH 8.1), 100 mM potassium acetate, and 30 mM magnesium acetate. Adjusted cRNA concentrations were used for fragmentation. The samples were incubated at 94°C for 35 minutes and put on ice following incubation.

2.3.4.11 Hybridisation of cRNA samples to Affymetrix Microarray Chips

Following fragmentation, 15ug of cRNA was mixed with a hybridisation cocktail containing eukaryotic and prokaryotic control RNAs used for validation of the efficiency of hybridisation. The target (the cRNA and hybridisation cocktail) was hybridised to the probe (Affymetrix MG-U74Av oligonucleotide microarray chips). 5 Dfl and 5 wild type targets each consisting of branchial arch dissections from 3 embryos were hybridised individually to 10 chips. In addition 4 microarrays were hybridised by Dr Sarah Ivins (MMU) using labelled cRNA from Elisabeth Lindsay, Baylor, Houston that was isolated in the same way from E10.5 wild type (n=2) and *Dfl* (n=2) embryos. The results from both the experiments were combined to improve statistical power.

For the *Splotch (Pax3)* enough cRNA (from branchial arch dissections from 3 pooled E10.5 embryos) was made for 3 wild type, 3 homozygote and 2 heterozygote GeneChip hybridisations.

Affymetrix MG-U74Av microarrays were prehybridized with 200µl 1x hybridisation buffer for 10 min at 45°C in an Affymetrix GeneChip Hybridisation Oven 640 at 60 rpm. Hybridisation was performed in a final volume of 300µl, containing 15µg fragmented biotinylated cRNA, 50 pmol control oligonucleotide B2 (Affymetrix), eukaryotic hybridisation controls (n =354; Affymetrix), 0.1 mg/mL herring sperm DNA, and 0.5 mg/mL of acetylated bovine serum albumin (BSA) in 1x hybridisation buffer. The samples were heated to 95°C for 5 min and 45°C for an additional 5 min and then briefly spun down. 200µl of the hybridisation cocktail was added to each GeneChip and hybridisations were performed for 16 hr at 45°C with mixing on a rotisserie at 60 rpm. After hybridisation, the solutions were removed, and the arrays were washed on a fluidics station (Affymetrix). Hybridized arrays were stained for 10 min at 25°C with streptavidin-R phycoerythrin (Molecular Probes; 10 µg/mL), followed by staining with biotinylated goat anti-streptavidin antibody (Sigma Chemical) then stained once again with streptavidin-R phycoerythrin for 10 min at 25°C.

Amplification of signal occurs by the initial staining with streptavidin-phycoerythrin which binds to the biotinylated nucleotides on the cRNA. This is followed by a second staining with biotinylated antibody which recognises the streptavidin epitope. The third stain was with streptavidin again which binds to the biotin conjugated to the antibody producing signal amplification. The phycoerythrin is then excited by a confocal argon-ion laser and emits fluorescent signals proportional to the amount of target bound to the arrays. The computer analysed the image file with a set target intensity of 100 and produced data files and report files. The report files were then examined to establish the quality of the hybridisation and arrays with poor quality were not used for data analysis.

Pixel intensities were measured, expression signals were analyzed and features extracted using a commercial software package (Microarray Suite version 5.0, Affymetrix). Data and statistical analyses were performed with Data Mining Tool Microarray Suite version 5.0, and Genespring version 4.2.1 (Silicon Genetics) bioinformatics algorithms.

2.3.5 Data Analysis using Affymetrix Microarray Suite

Each gene is represented by multiple 25-base oligonucleotides called probes, synthesized in square tiles on the surface of a silicon wafer. The probes come in pairs. One probe, the Perfect Match (PM), is designed to be complementary to a reference sequence. The second probe, or Mismatch (MM), controls for cross-hybridisation and contains a base pair mismatch at the central position (13^{th} bp). Typically each gene will be represented by 20 pairs of oligonucleotides referred to as probe sets. The surrogate expression level for a gene, its Average Difference, is measured by summing the differences between the PM and MM signals and dividing that sum by the number of probe pairs used in the calculation.

2.3.5.1 Detection Algorithm

The detection algorithm uses probe pair intensities to generate a detection p-value and assign a present, marginal or absent call. A discrimination score {R} is determined for each probe pair and describes its ability to detect its intended target. It measures the

target-specific intensity difference of the probe pair (PM-MM) relative to its overall hybridisation intensity (PM+MM):

$$\mathbf{R=(PM-MM)/(PM+MM)}$$

If the PM is much larger than the MM, the discrimination score for that probe pair will be close to 1.0. If the discrimination scores are close to 1.0 for the majority of the probe pairs, the calculated detection p-value will be lower and more significant. The detection p-value is then compared to the Tau threshold which was used at the default value of 0.015 in these experiments. The one-sided Wilcoxon's Signed Rank test is the statistical method employed to generate the detection p-value. It assigns each probe pair a rank based on how far the probe pair discrimination score is from Tau. This is illustrated in figure 2.3.5.1.

2.3.5.2 Detection call

Detection p-value cut offs provide boundaries for defining present, marginal or absent calls. Detection p-values below 0.04 determined a present call, between 0.04 and 0.06 determined a marginal call and between 1.00 and 0.06 determined an absent call. If a mismatch probe cell is saturated the probe pair is rejected from further analysis. In Genespring it is possible to identify a subset of genes below a specific detection p value cut-off. This filtering application was applied for some of my data sets (see specific chapters).

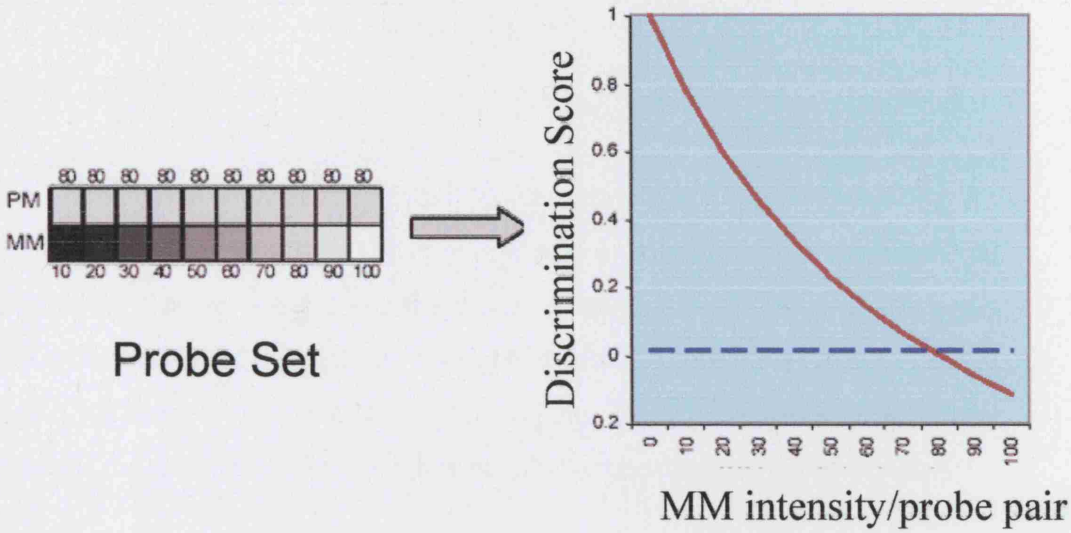


Figure 2.3.5.1

In this hypothetical probe set, the Perfect Match (PM) intensity is 80 and the Mismatch (MM) intensity for each probe pair increases from 10 to 100. The probe pairs are numbered from 1 to 10. As the Mismatch (MM) probe cell intensity, plotted on the x-axis, increases and becomes equal to or greater than the Perfect Match (PM) intensity, the Discrimination score decreases as plotted on the y-axis. More specifically, as the intensity of the Mismatch (MM) increases, the ability to discriminate between the PM and MM decreases. The dashed line is the user-definable parameter Tau (default = 0.015). Figure adapted from the Affymetrix User Manual.

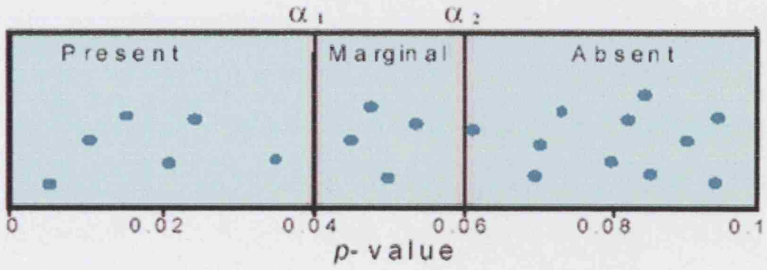


Figure 2.3.5.2

Significance levels α1 and α2 define cut-offs of p-values for detection calls. A lower detection p value signifies higher stringency and lower sensitivity and vice versa.
Figure adapted from the Affymetrix User Manual.

2.3.5.3 Signal Algorithm

Signal is a quantitative amount calculated for each probe set, which represents the relative expression of a transcript. Each probe pair in a probe set is considered as having a potential vote in determining the signal value. The signal intensity from the PM probes is used as an estimate of the real signal due to hybridisation of the target and the mismatch intensity is used to estimate stray signal. The real signal is estimated by taking the log if the perfect match intensity after subtracting the stray signal estimate.

2.3.6 Data Analysis using Genespring

Metrics files from the Affymetrix analysis files were loaded into Genespring version 4.2.1 (Silicon Genetics) for statistical analysis. The metrics files contained present/absent call information, detection p values and signal intensity data for each gene.

2.3.6.1 Normalisation of Raw Data

The following default normalizations appropriate to the raw data were applied by Genespring:

2.3.6.1.1 Per Chip Normalisation

The 50^{th} percentile of all measurements was used as a positive control for each sample; each measurement for each gene was divided by this synthetic positive control, assuming that this was at least 10. The bottom tenth percentile was used as a test for correct background subtraction. This was never less than the negative of the synthetic positive control.

2.3.6.1.2 Per Gene Normalisation

The measurement for each gene in each sample was divided by the corresponding value in the arithmetic mean of the control samples, assuming that it was at least 0.01. Lastly normalised values below 0.001 were set to 0.001.

2.3.6.2 Experiment interpretations

Two main methods were used to compare the wild type data set with mutant embryo data set. The first is the "mutant samples averaged" model where the raw intensity values were averaged for all of the samples and compared to the wild type data which were normalised to 1. In this model Genespring assumes all replicate samples in the same condition correspond to one kind of variability.

A second method of interpretation was the "mutant samples were treated as separate" where both wild type/mutant data and replicates were treated as continuous data. In this analysis each mutant data set was individually compared to the normalised wild type data set (still normalised to 1).

The interpretations applied to each dataset are explained in each chapter and illustrated in figure 4.3.5.1.

2.3.6.3 Gene filtering and statistical analysis

For each dataset the specific method of filtering and statistical analysis is described in each chapter.

Gene sets obtained from comparing the wild type and mutant data were filtered to obtain gene lists containing significantly changed genes. Initially genes with low control strengths were eliminated by inspecting the experimental data range setting and excluding genes below the "low control signal" cut off calculated by Genespring. The control strength is determined by the detection limit of a given gene and noise levels for any given measurement.

The genes were then filtered by fold change to include genes changing by a specific user defined cut off level. For instance values 0.5 or below corresponded to genes downregulated by at least 2 fold compared to the controls.

The statistical group comparison found genes with statistically significant differences in expression level between groups of samples. Statistical tests with applied to filter the genes. Several options are available at this point. For most of the analysis the parametric

test Welch t-test was used where the variances are not assumed to be equal. This was used with the multiple testing correction the Benjamini and Hochberg False Discovery Rate where the p values were ranked and a different threshold is applied for each. The smallest observed p value is compared against the strictest threshold, but the remaining p values are compared against successively more relaxed thresholds.

For the Df1 experiments I eliminated genes with low expression data by applying a detection p value cut off of 0.1 to the wild type dataset (i.e. mostly present and marginal calls included). This was done by using the data file restrictions option in the filtering and statistical analysis menu for the wild type group. The detection p-value was set to be less than 0.1 in at least 3 of the 7 replicates. This gene list was combined with the statistical analysis list in a Venn diagram and the group of genes in common was saved.

2.3.6.4 Cluster analysis

Experiment trees were created using the Genespring suite, following a per array normalisation using the distribution of all genes using the 50th centile, and a per gene normalization for each gene. Standard and Pearson's correlations were used to cluster arrays according to similarity. In order to identify a larger set of genes with changed expression patterns between genotypes, the Welch t-test was used to select genes with differing expression levels at various p values, without correction for multiple testing. This set (Cluster Set) of changed genes comprised 418 genes when $p<0.05$.

2.3.7 In situ hybridisation

2.3.7.1 Embryo dissection

Embryos (E10.5) were collected in calcium free PBS, dissected from their decidua and fixed in freshly made 4% formaldehyde diluted in PBS. They were left 4°C (sealed in a 48 well plate) overnight. The next day they were washed with 100% methanol, left for 5 minutes, covered in 100% methanol and put at –20°C sealed.

For each experiment 3-4 embryos were used each of wild type and mutant and the hybridisation reaction was done with all embryos in the same glass vials to ensure reaction conditions were identical. For each experiment a sense control was used as well as the antisense riboprobe to control for non specific binding.

2.3.7.2 Generation of labelled riboprobe

2.3.7.2.1 Template preparation

Depending on the orientation of the cloned cDNA, 10µg of template was linearised with 20units of an appropriate restriction enzyme in a total volume of 100µl at 37°C overnight. 1/20th reaction mix was checked on an agarose gel. DNA from the reaction mix was extracted with 100µl phenol (pH 7.0-8.0) :chloroform (1:1), vortexed and spun at 5000rpm for 2minutes and the top layer was removed into an eppendorph tube and the procedure repeated. The DNA was precipitated by adding 0.1 x volume of 3M sodium acetate (pH5.3) and 2.5 x volume of 100% ethanol (EtOH) mixed and placed at –70°C for 30mins or longer. The sample was spun for 15mins at 13,000rpm, the supernatant removed and the pellet was washed in 75% EtOH made up in DEPC treated water, vortexed and spun at 13,000rpm for 5mins. The pellet was air dried and resuspended in 1x DEPC H_2O at a concentration of 1µg/µl and stored at –20°C. The details of each riboprobe used for *in situ* hybridisation reactions are shown in table 2.3.7.2 below.

Table 2.3.7.2 Details of riboprobes made for *in situ* hybridisation reactions.

Gene	Affy no	Obtained from	IMAGE no	Vector	Enzymes	Primers
Connexin 45	104635	IMAGE consortium	495239	pCMV sport6	SalI/HindIII	T7/SP6
Rarg	102419	IMAGE consortium	3989967	pCMV sport6	SalI/HindIII	T7/SP6
FoxA2	93950	Deborah Henderson	n/a	pBLUEscript II KS+	BamHI/NotI	T7/T3
Tcf20	100947	IMAGE consortium	5368926	pCMV sport6	SalI/HindIII	T7/SP6
Tbx2	161146	Ginny Papaioannou	n/a	pBLUEscript II	BamHI/EcoRV	T7/T3
Sox3	103302	Robin Lovell-Badge	n/a	pBLUEscript II+	PstI/NotI	T7/T3
Wnt5a	99390	Deborah Henderson	99390	pgem7zf	XhoI/EcoRI	T7/SP6

2.3.7.2.2 *In vitro* transcription reaction

1μg of the linearised probe was transcribed with the appropriate enzyme (T7, T3 or SP6) for the vector in the following reaction:

Linearised template (1μg/μl)	1μl
MilliQ d H_2O	11μl
10x transcription buffer	2μl
DIG (Digoxigenin)-nucleotide mix	2μl
DTT (10X)	2μl
T7/T3/SP6 polymerase (20units/ul)	2μl
Total	20μl

Digoxigenin-nucleotide mix consisted of 10mM GTP, 10mM ATP, 10mM CTP, 6.5mM UTP, 3.5mM Digoxigenin-labelled UTP.

The transcription mix was incubated at 37°C for 2hr and 1μl of the probe was electrophoresed on 0.8% agarose gel to check the reaction had worked.

2.3.7.3 EtOH precipitation of probe

2μl 0.2M EDTA, 2.5μl 4M LiCl and 75μl 100% EtOH (-20°C) were added to the reaction and mixed well. The probe was placed at –70°C for 30mins or overnight, spun at 12,000rpm for 15mins at 4°C and the pellet was rinsed in 300μl 70% EtOH in DEPC H_2O, vortexed and spun at 12,000rpm for 5mins and dried in air. The pellet was resuspended in 90μl DEPC H_2O (final concentration is approximately 1mg/ml as transcription should yield about 8x the original weight of DNA) and stored at –80°C.

2.3.7.4 Prehybridisation

Embryos were put in 5 or 10ml glass vials and rehydrated with decreasing concentrations of methanol (MeOH). They were then washed three times with PTw (PTw: PBS, 0.1% Tween-20) for 5 minutes with horizontal rocking. Proteinase K was added to make a final concentration of 10μg/ml and incubated at room temperature for

20 minutes. The proteinase K was removed and the embryos were carefully rinsed with PTw. The PTw was replaced with 4% paraformaldehyde in PTw containing 0.1% gluteraldehyde. They were then fixed for 20-30 mins with horizontal rocking.

The postfixing solution was removed and the embryos were and washed with PTw three times for 5 minutes at room temperature. All but 1ml of PTw was removed and 1ml of hybridisation solution was added without probe (see below for composition). This was replaced with 1ml of hybridisation solution without probe and left to prehybridise at 68°C with vertical shaking overnight.

2.3.7.5 Hybridisation with riboprobe

The embryo containing vials were removed from the incubator and placed at -20°C. 6 hours later the hybridisation mixture with the riboprobe was added to the embryos at a final concentration of 1µg/µl and incubated at 68°C overnight.

2.3.7.6 Post hybridisation washes

The riboprobe was removed and the embryos were rinsed twice with prewarmed hybridisation mix. They were washed three times with prewarmed hybridisation solution at left at 68°C in for 30 min. They were then washed for 20 minutes with prewarmed hybridisation buffer:TBST (1:1) which consists of 2.5ml hybridisation solution, 0.25ml 10X TBST, 2.25ml H_2O and 0.1% tween. The embryos were rinsed with TBST and washed three times at 30 minutes each with TBST. The embryos were blocked with 20% sheep serum in TBST for 2-3 hours at room temperature. The blocking buffer was removed and 1/2000 anti-DIG antibody was added (0.5µl antibody in 1000µl blocking buffer). The embryos were incubated overnight at 4°C on a rocking platform.

2.3.7.7 Post-antibody washes

The antibody solution was removed and the embryos were rinsed three times with TBST. They were then washed twice for 10min a time with NTMT. NTMT contains 4.5µl NBT (Nitroblue tetrazolium stock 75mg/ml in 70% dimethylformamide) and 3.5µl BCIP (5-bromo-4-chloro-3-indolyl-phosphate stock 50mg/ml in 100% dimethylformamide) per 1.5ml. The embryos were left rocking on a platform protected

from light until they went blue. After colour had developed the reaction was stopped by washing twice 10min in TBST and fixed in 0.1% formaldehyde and stored at 4°C.

2.3.8 Real Time Quantitative PCR

2.3.8.1 RTQPCR analysis of Df1 and wild type embryos

For validation of the microarray results, total RNA was extracted from the branchial arch region of E10.5 mouse embryos (see figure 2.3.4.2) using TRIzol. Branchial arch dissections were performed on embryos from 18-36 somites and embryos were genotyped using yolk sac DNA. Embryos were separated into *Df1/+* and wild type and pools of between 3 and 6 embryos were made for 9 different somite stages. Analysis for all experiments except the *FoxA2/Shh/Tbx1* comparison were performed with cDNA from the 30, 32, 34 and 36 somite stages in triplicates. Experiments were performed for at least three of these four somite stages (with the exception of *Sox11*, *Sox3* and *Smarca5* where experiments were performed for two somite stages due to time constraints). Relative ratios were averaged across these stages to give the final relative change ratios in the results section. For the *FoxA2/Shh/Tbx1* comparison experiment (section 5.3.1.8) cDNA was used from embryo batches at 18, 22, 28 and 34 somites and the experiment was performed once in triplicate.

2.3.8.2 RTQPCR analysis of Tbx1 and wild type embryos

Embryonic day 10.5 mouse embryos from *Tbx1/+* vs wild type crosses were dissected and RNA was extracted as described. The embryos were genotyped using yolk sac DNA and pooled into 4 embryos each of *Tbx1* heterozygotes and wild types.

2.3.8.3 RTQPCR using Pax3 heterozygous, homozygous and wild type embryos

For *Pax3* work embryos were more limited in supply. 3-4 embryos were combined for each genotype (wild type, homozygotes, heterozygotes) at the 30-34 somite stage. RTQPCR experiments were performed at least three times for each batch and relative ratios averaged.

RNA was extracted using TRIzol and was then purified using the RNeasy clean-up kit (QIAgen) and treated with RNase-free DNase (Promega) to eliminate genomic double-stranded DNA which would interfere with the RTQPCR reaction.. The concentration was calculated using UV spectrophotometry and quality of the RNA assessed using 1% gel electrophoresis. RNA was discarded if either the yield or quality were below expected levels. 2μg of pooled wild type or mutant RNA was transcribed using Superscript II reverse transcriptase and random hexamer primers (Invitrogen) according to the manufacturer's instructions, in a total volume of 20μl. As a control 2μg of wild type tRNA was used without reverse transcriptase. The resulting cDNA was diluted to use in a quantitative PCR reaction. The dilution factor was variable and each cDNA was tested to ensure CT values were not higher than 30 as accuracy is compromised beyond this level of dilution RTQPCR was performed using SYBRGreen (QIAgen) for the detection of fluorescence during amplification. Each amplification reaction contained of 0.1μg of forward and reverse primers, 6.25μl of SYBRGreen mix and 5.25μl of diluted cDNA in a total volume of 12.5μl. All primers were designed with annealing temperatures of 57-61°C.

GAPDH was used as an internal control as there was no difference in *GAPDH* expression levels between wild type and *Df1* cDNA on the microarrays. PCR conditions were 95°C for 15mins, followed by 40 cycles of 95°C for 30 secs, 57-61°C for 30 secs, and 72°C for 30 secs. All reactions were performed in triplicate in a DNA Engine Opticon 2 system (M. J. Research inc.) and a threshold cycle (Ct) value of 0.001 was used for analyses. The relative expression ratios were arrived at by using the $2^{-\Delta\Delta Ct}$ method of data analyses (Livak & Schmittgen 2001) with *GAPDH* levels used to standardise the cDNA input quantities. The exact details of this method are explained in chapter 5 with (figure 5.3.1.3). Details of the primer sequences and annealing temperatures for the genes tested are shown in table 2.1.7.4.

2.3.9 Fluorescence in situ hybridisation

2.3.9.1 Preparation of metaphase spreads

Metaphase spreads were prepared from healthy cell suspensions in a fume hood. Cells had been incubated with 250 μl colcemid (Sigma; 10 μg/ml) for 1 hour at 37 °C in order to induce mitotic arrest. Cells were harvested by centrifugation at 1000 rpm for 5 minutes and most of the medium was aspirated. The pellet was resuspended gently in

the remaining medium and 10 ml hypotonic 75 mM KCl was added dropwise to form an even suspension. Cells were immediately re-centrifuged at 1000 rpm for 5 minutes and most of the supernatant was aspirated. 2 ml of fixative (methanol: acetic acid; 3:1) was added dropwise so that the cells were quickly and thoroughly resuspended. Cells were centrifuged again as above, the supernatant was removed and the cells finally resuspended in a few drops of fixative until the solution was just translucent. Superfrost slides were prepared by soaking in absolute ethanol followed by polishing with lint-free tissue. The slides were breathed on to humidify the slide. The cell suspension was dropped onto the prepared slides from 4cm and a drop of fixative was added as the centre of the drop started to dry. The spread was examined under the microscope using phase contrast. The area to be hybridised was marked with a diamond tipped pencil. Slides were stored in the short term at 4 °C in a sealed box.

2.3.9.2 Preparation of BAC probes

Appropriate Bacterial Artificial Chromosomes (BACs) were chosen from BACPAC (http://www.chori.org/bacpac/) and DNA was isolated using the QIAgen Midi Prep protocol for low copy number plasmids. The final DNA pellet was dissolved in 100-200µl TE buffer. Quality and quantity was determined by spectroscopy.

2.3.9.3 Nick translation probe labelling

The BAC DNA was labelled with SpectrumRed-dUTP or SpectrumGreen-dUTP (Vysis Inc, Illinois, USA) by nick translation (Vysis). This method generates labelled 200-500 bp fragments suitable for FISH. The kit includes an enzyme mix of DNase I, which nicks the DNA, and *E. coli* DNA polymerase I, which generates the nick translation fragments. The labelling reaction was set up on ice to include 1 µg purified BAC DNA, 2.5 µl 0.2mM SpectrumGreen, SpectrumOrange or SpetrumRed dUTP, 10 µl 0.1mM dNTP mix (10µl each of 0.3mM dATP, 0.3mM dCTP and 0.3mM dGTP), 5µl 0.1mM dTTP, 10 µl nick translation enzyme, 5µl 10X nick translation buffer in a total volume of 50 µl made up with ddH_2O. The reaction was incubated at 15 °C for 15 hour and then stopped by incubation at 70°C for 10 minutes and then placed on ice. A 5 µl aliquot was electrophoresed on a 1 % agarose gel to check the fragment size and efficiency of the nick translation.

2.3.9.4 Hybridisation, washes and detection

The labelled probe was mixed with 2 µl Cot 1 (1mg/ml) human competitor DNA (50-350 bp fragments of human placental DNA, enriched for repetitive sequences) and 0.4 µl herring testes DNA (10mg/ml) in order to block repetitive sequences. The DNA was then co-precipitated by the addition of 2µl of 3 M Na acetate (pH 5.2) and 56µl of cold absolute ethanol and 8µl ddH_2O followed by incubation at –80 °C for 30 minutes. After centrifugation at 12,000rpm for 30 minutes at 4°C, the supernatant was aspirated and the DNA pellet was rinsed with 70% ethanol, air-dried and resuspended in 20 µl hybridisation solution (50 % deionised formamide, 10 % dextran sulphate, 2× SSC), left at 37°C overnight and put at -20°C the next morning.

Freshly prepared slides of metaphase spreads were aged by immersion in a Coplin jar containing 2× SSC at 37 °C for 1 hour, then dehydrated by passage through a series of 3 washes of 70 %, 90 % and 100 % ethanol for 5 minutes each, at room temperature.

For hybridisation, 5µl of hybridisation solution containing the labelled BAC probe and 5µl of each control probes {5p13.3 Cri du Chat probe (D5S721) and 5q EGR-1 (D5S23)} was pipetted onto each slide. An 18 × 18 mm coverslip was placed over the slide, any bubbles were expelled and the coverslip was sealed with weldtite vulcanising rubber solution. The slides were then denatured at 73°C for 1 minute and incubated overnight at 37 °C in a wet chamber.

After hybridisation, the coverslips were removed and any unbound or weakly bound probe was removed by washing the slides in 2X SSC and Tween-20 (ST). The slides were then put in a Coplin jar prewarmed with 0.4%SSC/0.3%NonidetP-40 (BDH) at 73°C for 2 minutes. They were then washed in ST for 2 minutes and allowed to thoroughly dry.

10µl of Vectashield with DAPI was applied to each slide and overlaid with a 22mm X 22mm coverslip. DAPI (4,6-diamino-2-phenyl-indole) is a fluorescent dye, which produces a banding pattern on the chromosomes enabling them to be visualised. Vectashield is a commercial mount, which helps to prevent photo-bleaching. The slides

were stored in a light proof box at 4 °C. Excess DAPI was blotted away using blotting paper and a thin layer of glue was applied to hold the coverslip in place.

Metaphase spreads were visualised using a Leica DMLB fluorescence microscope (Leica Microsystems Wetzlar GmbH) and images were captured using a MetaSystems ISIS workstation MetaSystems GmbH, Altlussheim).

2.3.10 Bioinfomatics

2.3.10.1 Database homology searches

In order to investigate possible matches between a query sequence and a nucleic acid or protein sequence in the database, BLAST (Basic Local Alignment Search Tool) was used at http:/www.ncbi.nih.gov/BLAST (Altschul et al. 1990). "Two way BLAST" analysis was used to examine homology between two known sequences used for probe design analysis in Chapter 4.

2.3.10.2 Sequence alignments

Alignments of two or more protein or nucleic acid sequences was performed using programs available at the SDSC Biology Workbench at http://workbench.sdsc.edu/.

Chapter 3 Yeast Two Hybrid analysis of Tbx1

3.1 Introduction

Protein-protein interactions are essential for most biological processes, and their study is important in understanding cellular mechanisms. Since its introduction, the yeast two hybrid system has been widely used to identify interactions between proteins in eukaryotes (Fields & Song 1989). The system can be used to test for a direct interaction between two known proteins, or to identify previously unknown interacting proteins by utilising cDNA expression library screening. The yeast two hybrid system was initially tested using two yeast proteins, SNF1 and SNF4 which are known to interact (Fields & Song 1989). The system can also be utilised to localise the protein domains responsible once an interaction has been identified.

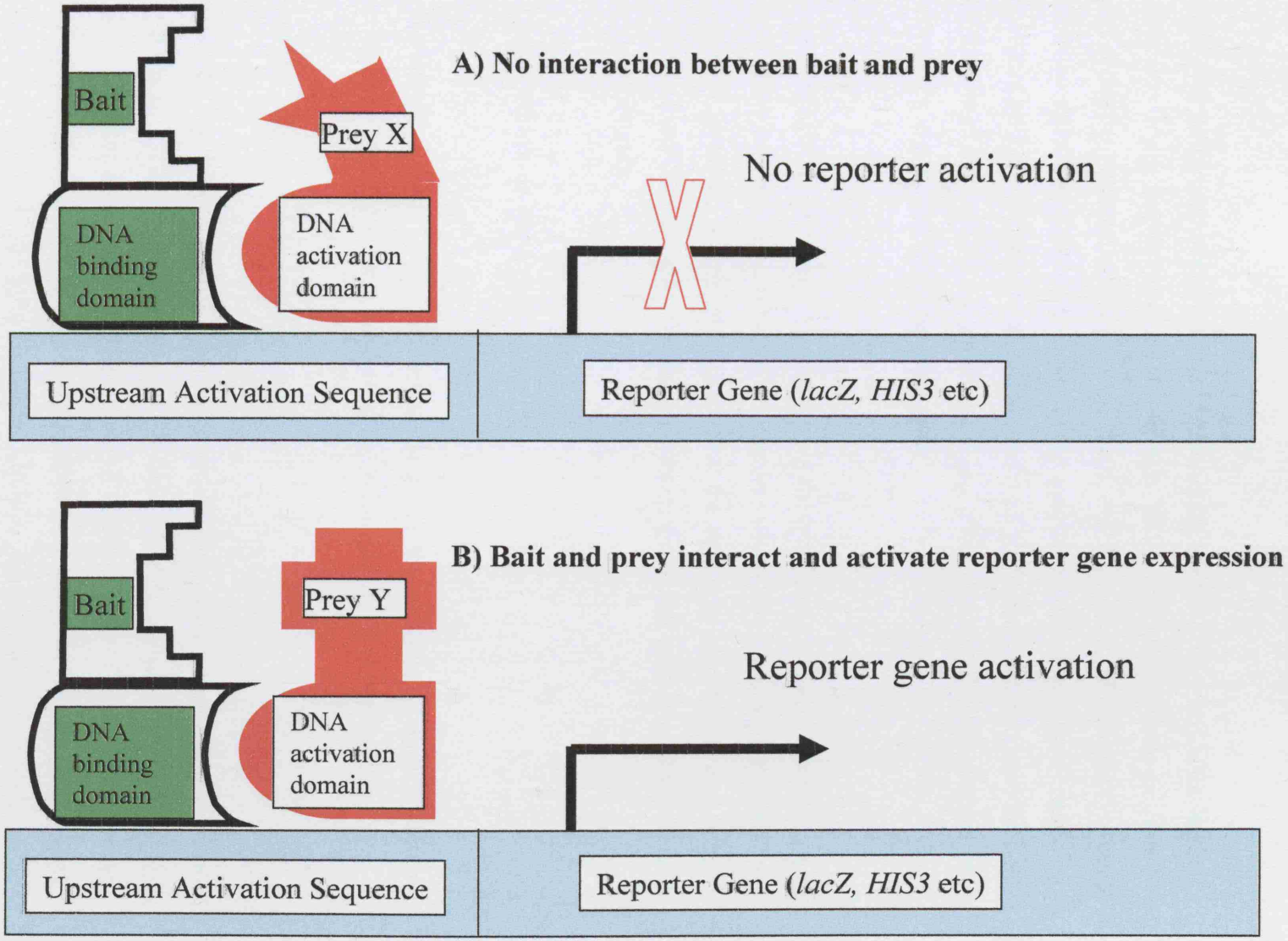
3.1.1 The Yeast Two Hybrid System

The yeast two hybrid system takes advantage of the fact that certain eukaryotic transcription factors are comprised of two distinct and separable domains: a specific DNA-binding domain (DBD) and an activation domain (AD), each functionally essential. The protein of interest is designated the "bait" protein and is encoded as a cDNA fusion to a DBD protein, for example, the GAL4 transcriptional activator of *Saccharomyces cerevisiae* (required for the expression of genes encoding enzymes of galactose utilisation). The DBD binds to specific upstream activating sequences (UAS) within yeast promoters and transcription is only activated when the DBD associates with the acidic regions of the AD. The "prey" proteins (either a candidate protein interactor or a cDNA library) are fused to the transcriptional AD (figure 3.1.1).

The system involves the co-transformation of two plasmids containing the gene fusions into a yeast host strain containing inducible reporter genes. The individual plasmids cannot activate transcription alone and require a protein-protein interaction between the bait and the prey to bring the DBD and the AD into close proximity to activate transcription of the reporter genes. If there is no interaction between the bait and prey, the DBD and AD remain separate and transcription is not initiated. When novel interactions are being sought, an activation domain library prepared from a specific source of cDNA is used to generate the prey.

Figure 3.1.1 The Yeast Two Hybrid System.

There are two separate domains to a eukaryotic transcription factor, the DNA binding domain which binds to a specific site in the upstream activation sequence and the activation domain. The two domains of the transcription factor are fused to the bait (the protein if interest) and the prey (in the form of a library of candidate proteins). A protein-protein interaction between the bait and the prey is required to bring the DNA binding domain and the activation domain into close proximity to activate transcription of the reporter genes. A) illustrates the situation where the prey X does not interact with the bait and B) illustrates reporter gene activation when there is an interaction between the prey Y and the bait.



The most frequently used transcription factor modules are the GAL4 or LEXA binding domains, the GAL4 activation domain, and the VP16 activation domain of *Herpes simplex*. A commonly used reporter gene is the *E. coli lacZ* gene, which hydrolyses the chromogenic substrate X-Gal, producing blue colonies on plates or filters containing this indicator. In addition, nutritional markers such as genes encoding enzymes involved in biosynthesis of amino acids e.g. *HIS3* and *ADE2* allow selection for colonies that can grow on media lacking the relevant amino acid.

The yeast two hybrid system is very sensitive as the hybrid proteins are over-expressed by strong promoters on high-copy-number vectors, which favours complex formation. The system is therefore somewhat artificial and the resulting proteins isolated from a library screen must be viewed as 'candidates'. Interactions will require confirmation by independent methods.

3.1.2 Limitations of the system

The two hybrid system is an efficient method for the analysis of protein-protein interactions. However there are a number of limitations that need to be considered.

3.1.2.1 False Positive Protein interactions

False positive interactions are colonies that grow when the reporter gene is transcribed in the absence of an interaction between the bait and prey. There are a number of possible reasons for this including the problem of autoactivation. This occurs when the protein fused to the GAL4 DBD induces activation of the reporter constructs alone, without interaction with the prey construct. 5-10% of proteins, particularly those containing a high proportion of acidic residues, are able to induce transcriptional activation when fused to a DBD (Toby & Golemis 2001). In order to identify this class of false positive, the bait alone must be transformed into the host yeast strain prior to a library screen, to see whether any of the reporter genes are activated. If autoactivation does occur it is necessary to map the autoactivation domain of the bait protein and use fragments of the protein that do not contain this region, although this has the possible disadvantage of affecting the protein's interactions.

The second problem is the interaction of the GAL4 protein with an endogenous yeast protein resulting in reporter gene activation. A method for reducing this problem is to use two or three distinct promoter constructs for the reporter genes. Also use of a non yeast DBD protein such as the *E. coli* transcription factor LEXA results in a lower chance of cross-reaction with yeast proteins.

Lists of commonly occurring false positives that have been detected in yeast two hybrid screens using a variety of different proteins are available online:
www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html

3.1.2.2 False Negative Protein interactions

These can occur in the system for several reasons. For example the introduction of some hybrid bait proteins appears to interfere with normal cellular events in yeast to such an extent that the cells are non-viable (e.g. apoptosis inducing genes). The problem of baits which exhibit such toxicity may be overcome by truncation of the bait, but again this can have the disadvantage of affecting the protein's interactions.

In addition, the conformation of the protein may be altered significantly by the presence of the activation or binding domain, such that interactions are altered due to steric hindrance. False negatives can also occur when interactions that are dependent on post-translational modifications of either bait or prey taking place on the endoplasmic reticulum, are not detected.

However, despite these limitations, the yeast two hybrid system can be a very powerful tool with which to begin the functional characterisation of a protein.

3.2 Experimental Methodology

Full details or the methodology used for this yeast two hybrid screen are available in section 2.3.3 (Chapter 2).

3.2.1 Choice of Yeast strain PJ69-4A

PJ69-4A is deficient in *TRP1*, *LEU2*, *URA3*, *HIS3*, *ADE2* and *MET2* genes and is only able to grow if tryptophan, uracil, histidine, adenine and methionine are added to the growth medium (James et al. 1996). This growth restriction allows the future selection of recombinant interacting clones.

There are other modifications that are necessary to ensure the yeast strain is useful in this system. The yeast chromosomal copies of the *GAL4* and *GAL80* genes must be deleted. Endogenous expression of *GAL4* would result in constitutive expression of the reporter gene regardless of whether there is an interaction between the bait and the prey proteins. In addition, as *GAL80* is a repressor of *GAL4*, the gene must be deleted to allow transcription of the reporter genes. If a gene is to be used as a reporter gene to identify positive colonies in which the bait and prey have interacted, the native promoter must be removed and replaced with a promoter containing an UAS.

The PJ69-4A yeast strain has three different inducible reporter genes, *Gal2-ADE2, Gal1-HIS3* and *Gal7-lacZ*, each driven by a different GAL4-responsive promoter. Selection using the *Gal2-ADE2* is more stringent than the use of *Gal1-HIS3* and is thus used as the second selection in these experiments. The *ADE2* gene confers a colour phenotype such that colonies which are synthesising adenine via an alternative pathway rather than through the reporter gene can be identified by a pink or red colour of the colonies. Thus colonies of this colour phenotype were rejected as false positives. X-α-Gal was incorporated into the plates lacking adenine to test for activation of the *Gal7-lacZ* promoter by a bait-prey interaction.

As *Gal1-HIS3* construct is more stringent than in most other strains, PJ69-4A has the advantage of requiring very low concentrations of 3-aminotriazole (3-AT). This is a competitive inhibitor of the *HIS3* gene product, used to eliminate background *HIS3* expression and further refine the histidine selection for proteins that show a more

pronounced interaction. At high concentrations, 3-AT inhibits cell growth, but at the lower concentrations required by PJ69-4A, typically 0-2 mM, cell growth is not inhibited.

3.2.2 Choice of cDNA library

The choice of library or "prey" for a yeast two hybrid screen is important as the activation domain fusion libraries can be prepared from different organisms, tissues, cell types and developmental stages. A whole mouse embryo E9.5 and E10.5 random primed cDNA library in pVP16 was used because this resource was available and had previously been used successfully in the laboratory. *Tbx1* is strongly expressed at this time point in mouse embryogenesis, therefore establishment of potential protein binding partners at this dynamic time just before arterial remodelling was considered an advantage.

The pVP16 cDNA library vector (Appendix 1) utilises the VP16 activation domain of *H. simplex* driven by the ADH promoter and carries a leucine marker gene, *LEU2*. The mouse embryo cDNA inserts are cloned into the *NotI* site.

3.2.3 Cloning of GAL4 fusion products for Yeast Two Hybrid analysis

Three constructs were designed to represent the N terminus (nucleotides 4-333), C terminus (nucleotides 825-1467), and full length (nucleotides 4-1467) mouse Tbx1. The T-box site (DNA binding domain) was not included in the C and N terminus constructs. These coding sequences were cloned in frame with a cassette encoding the GAL4 DNA binding domain in order to generate mTbx1-GAL4 fusion proteins.

3.2.4 Identification of positive clones

Prey plasmids were extracted from positive clones as described in 2.3.3.7. The insert was amplified using VP16 primers and the PCR product sequenced. The identity of the clone was determined using BLAST (www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). Following identification of the clone, the reading frame was determined to ensure that the cDNA was in frame with the VP16 activation domain. The orientation of the clone and identification of whether the clone's homology was within the coding or noncoding region of the cDNA was established.

3.3 Results

3.3.1 Preliminary experiments

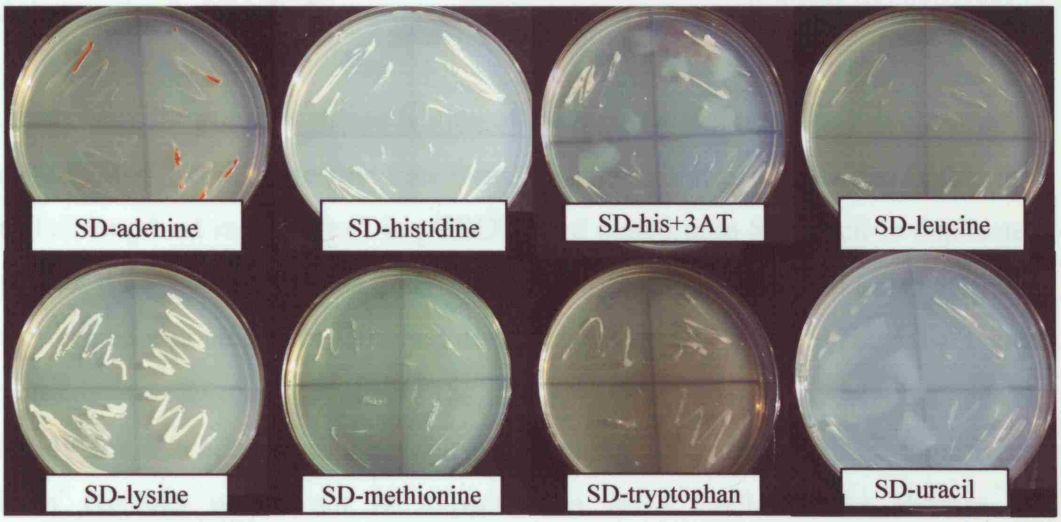
3.3.1.1. Confirmation of PJ69-4A phenotype

It is important to confirm the phenotype of the yeast strain prior to the commencement of each yeast two hybrid experiment as the cells have the potential for mutation while being recovered from glycerol stocks. In order to test the growth requirements of the yeast strain, the PJ69-4a was plated onto SD media with each of the amino acids missing in turn and left for 48 hours at 30°C. The results are shown in figure 3.3.1.1. It can be seen that there is no, or minimal, growth for the plates lacking leucine, methionine, tryptophan and uracil and the plate lacking histidine with 2mM 3-AT added. There was growth on the SD-histidine plates as PJ69-4A has some background expression of the *HIS3* gene product. To overcome this 2mM 3-aminotriazole (3-AT) was added to the plates. It is a competitive inhibitor of the *HIS3* gene product and eliminates background *HIS3* expression. The plate lacking adenine has minimal growth with a red tinge which indicates the colonies are synthesising adenine via an alternative pathway. Positive growth on the SD-lysine plate served as a positive control. The phenotype of PJ69-4A was as expected and was suitable for use in the yeast two hybrid system.

3.3.1.2 Sequencing of Constructs

All three constructs were sequenced and were found to be in frame with the GAL4 cDNA and the C and the N terminal constructs had no mutations induced by PCR. The full length construct was found to have 3 missense mutations, but these were not causing an amino acid change.

Figure 3.3.1.1 Genotype analysis of PJ69-4a. Note growth on SD-histidine and SD-lysine which is expected.



Example of a colony activating the *LacZ* reporter gene.

Example of a colony not activating the *LacZ* reporter gene.

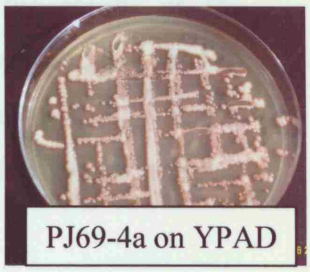
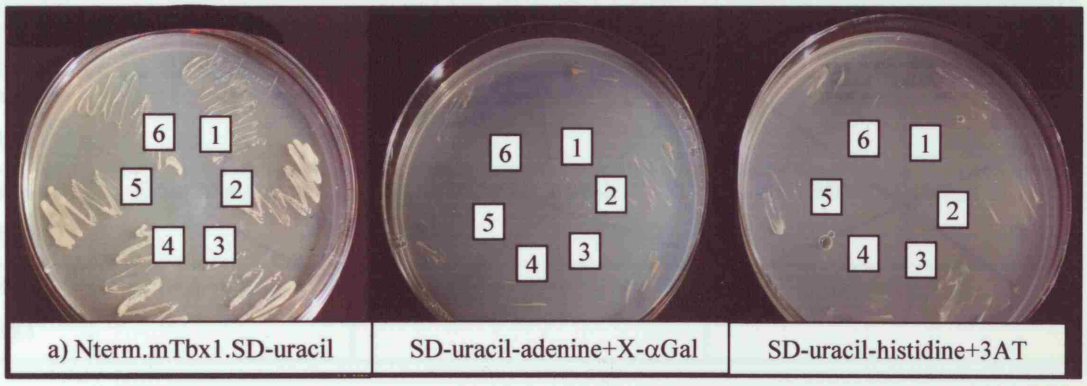
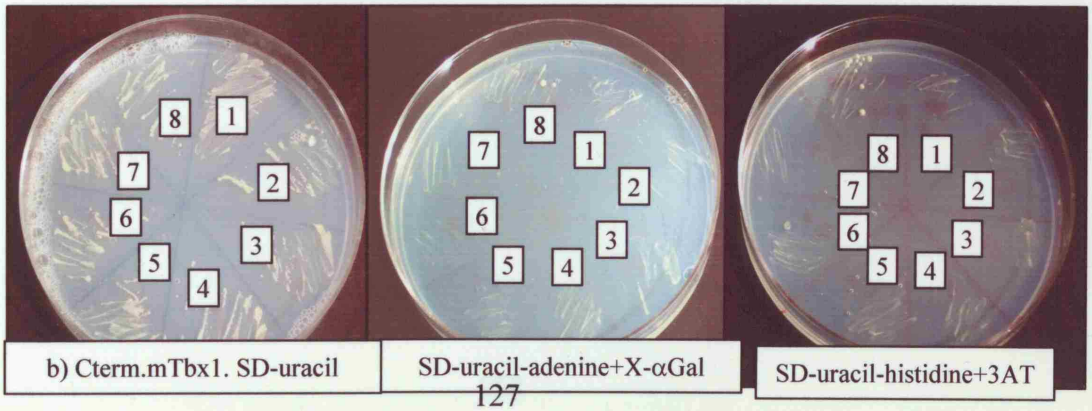


Figure 3.3.1.2 Test for autoactivation of a) N term.mTbx1-GAL4 fusion construct and b) C term.m.Tbx1-GAL4 fusion construct. Note lack of growth on SD-uracil-adenine +X-αGal and SD-histidine +3AT indicating no autoactivation for the N and C terminus constructs.





3.3.1.3 Autoactivation analysis

In order to check for possible autoactivation of the reporter genes by the bait constructs, competent PJ69-4A cells were first transformed with the bait plasmid alone (either the C, N or full length of mouse Tbx1 in pGBDU) and plated onto SD-uracil in duplicate. As expected, the transformed cells grew successfully on SD-uracil by 48 hours at 30°C. The transformed colonies were then replated onto SD-uracil-histidine+2mM 3-AT and SD-uracil-adenine+X-αGal to check the *HIS3*, *ADE2,* and *LacZ* reporter genes respectively for autoactivation. They were grown for 72 hours at 30°C. Significant growth of colonies would not be expected on these media in the absence of autoactivation. Figure 3.3.1.2a and b and table 3.3.1.2 show the results of the autoactivation screen. The N and C terminal constructs did not autoactivate, but the full length mouse Tbx1 did on the SD–histidine+2mM 3AT medium and was inconclusive on the SD-adenine+X-αGal medium (all 6 colonies were red). The full length construct was therefore not used for the yeast two-hybrid screen.

Table 3.3.1.3

Summary of autoactivation analysis with full-length and truncated baits after 72 hours of incubation at 30°C.

GAL4 fusion construct	SD-uracil	SD-uracil–adenine+X-αGal	SD-uracil-histidine+2mM 3-AT
N terminus mTbx1	Extensive growth	Minimal growth	Minimal growth
C terminus mTbx1	Extensive growth	Minimal growth	Minimal growth
Full length mTbx1	Extensive growth	Colonies appeared red	Extensive growth

3.3.1.4 Western analysis

An analysis of expression of the yeast two hybrid constructs was done by the use of a Western blot with an antibody to the GAL4 protein. The results of the Western analysis are shown in figure 3.3.1.4. The approximate weight of the protein is determined by the following equation:

Number of amino acids X 110= approximate weight in Daltons.

The N terminal region is calculated to be 12kDa and the C terminal region 20kDa. The GAL4 binding domain is approximately 16 kDa. Thus on Western analysis the bands for the N and the C terminal-GAL4 fusion constructs were expected to appear at the 28kDa and 36kDa marks respectively. The C terminal GAL4 fusion protein was detectable by Western analysis while the N terminal fusion construct was not. It was decided to proceed with the yeast two hybrid protocol for both of these constructs despite lack of evidence for expression of the N terminus, as experience in the lab has shown that the experiment can still work without detection on a Western blot. This is because expression can sometimes be transient or too low for the sensitivity of the system. Western analysis was not performed for the full length *mTbx1* construct as it was found to autoactivate.

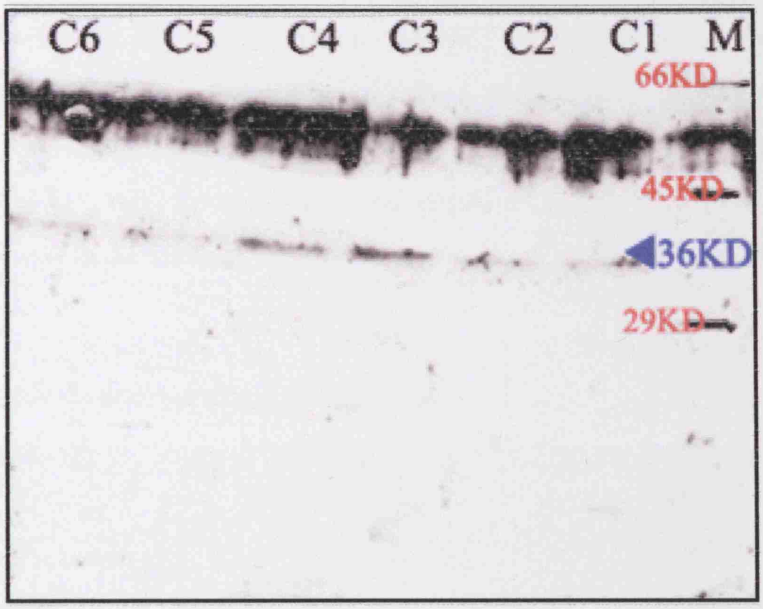


Figure 3.3.1.4

Western analysis showing detection of the mouse Tbx1 C Terminal-GAL4 fusion construct using an anti-GAL4 antibody. C1-C6 indicate protein aliquots from different colonies of PJ69-4A transformed with the construct. M denotes the protein marker. The protein is expected to be about 36 kDa (KiloDaltons) as indicated by the marker on the right of the autoradiograph. The N terminus was not detected by Western analysis.

3.3.1.5 Results of Transformations

Table 3.3.1.5 summarises the results of the two yeast two hybrid screens performed. Both constructs were initially transformed at one tenth the concentration of the full scale transformation. The C terminal yeast two hybrid screen gave approximately 1500 colonies and N terminus screen gave six. It was therefore decided to only proceed with the full scale transformation for the N terminal construct and to use the one tenth scale results for the C terminal construct.

Table 3.3.1.5

Results of yeast two hybrid transformations.

	N terminal *mtbx1* **(Full Scale transformation)**	**C terminal *mtbx1*** **(Small scale: 1/10 of full scale transformation)**
Transformation Efficiency	1.5×10^8 clones using 250µg of "bait" DNA	2.3×10^7 clones with 25µg of "bait" DNA
No. colonies on SD-uracil-leucine-histidine+3AT selection	67	Approximately 1500 colonies grew-512 put on selection plates
No. colonies on SD-uracil-leucine-adenine selection.	30	395
No. blue colonies with X-α-Gal.	18	357
No. colonies with successful plasmid PCR.	16 (including 1 clone with 2 plasmids)	28 (including 3 with 2 plasmids)
No. colonies successfully sequenced.	16	30

3.3.1.6 Potential interactors of Yeast Two Hybrid screen: C terminal construct

Table 3.3.1.6 shows the list potential protein interactors identified through the yeast two hybrid screen using the C terminus of mouse Tbx1 as bait. 14 of the 30 plasmids gave no significant homology to any proteins on NCBI BLAST analysis. All of the clones identified were either homologous to non coding regions (n=6), not in frame (n=1), not in the correct orientation (n=1), of unknown significance (n=4) or likely false positives (n=4) by virtue or their cellular location.

3.3.1.7 Potential interactors of Yeast Two Hybrid screen: N terminal construct

Table 3.3.1.7 shows the list potential protein interactors identified through the yeast two hybrid screen using the N terminus of mouse Tbx1 as bait. Five of the 16 plasmids showed significant homology to two separate proteins on BLAST analysis. Both of the potential interactors were in frame in the correct orientation and homologous to the coding region of the gene. The first protein was mouse peroxiredoxin 4 which is an antioxidant enzyme and a likely false positive.

The second protein identified was the cytoplasmic dynein light chain 1 (DLC1: NM_019682, Q15701) and sequence matching resulted in the identification of four independent clones. DLC1 consists of 531 nucleotides with a coding sequence from nucleotides 100-369. The protein is 89 amino acids in length and is 8kD. There were seven colonies of clone A which spanned from nucleotides 65 to 424 encompassing the entire open reading frame (amino acids 1-89). Clone B was identified from two colonies and spanned from nucleotides 3 to 320 (amino acids 1-73). Clone C was also sequenced from two colonies and spanned from nucleotides 58 to 470 (amino acids 1-89) and clone D was found in one colony and spanned from nucleotides 28-470 (amino acids 1-89). DLC1 is not listed among the false positives in a database of two-hybrid artefacts (http://www.fccc.edu/research/labs/golemis/Table1.html), indicating that DLC1 is not known to interact randomly with a large number of related proteins. Also this interactor has not been reported in yeast two hybrid screens from any other groups in the laboratory that have used the same library.

	Potential Interacting protein	Orientation	Coding/ Noncoding	No of Copies.	In Frame	Interpretation (see text)
1	Mus musculus proteasome activator gene	Correct	Non	1	N/A	Non coding
2	Mus musculus jagged 1 (Jag1)	Correct	Coding	1	Yes	Probable false positive
3	Nesp55 gene for neuroendocrine secretory protein 55	Correct	5'UTR	3	N/A	Non coding
4	Similar to polypyrimidine tract binding protein	Correct	5'UTR	1	N/A	Non coding
5	Mouse DNA sequence from clone RP23-37J21	Correct	Coding	2	N/A	Unknown significance
6	Mus musculus similar to Kielin (LOC243735),	Correct	Coding	1	Yes	Probable false positive
7	Homo sapiens chromosome 19 clone RP11-886P16	Correct	Coding	1	N/A	Unknown significance
8	Mus musculus, poly A binding protein, cytoplasmic 1	Correct	Non	1	N/A	Non coding
9	Mus musculus melanoma antigen, family D, 2 (Maged2), mRNA	Correct	N/A	1	No	Incorrect Frame
10	Mus musculus LOC240312 (LOC240312), mRNA	Correct	Coding	1	N/A	Unknown significance
11	Similar to MEGF10 protein-EGF-like domains-extracellular	Incorrect	No	1	N/A	Incorrect orientation
12	Mus musculus granulin (Grn)	Correct	Coding	2	Yes	Probable false positive

Table 3.3.1.6 Results of Yeast Two Hybrid screen using Cterminus mTbx1 as bait.

	Potential Interacting Protein	Orientation	Coding/ Noncoding	No of Copies.	In Frame	Interpretation (see text)	Homology (nucleotides)
1	Mus musculus, peroxiredoxin 4	Correct	Coding	1	Yes	Probable false positive	
2A	Mus musculus, dynein, cytoplasmic, light chain 1	Correct	Coding	7	Yes	Possible true positive	65-425
2B	Mus musculus, dynein, cytoplasmic, light chain 1	Correct	Coding	2	Yes	Possible true positive	3-320
2C	Mus musculus, dynein, cytoplasmic, light chain 1	Correct	Coding	2	Yes	Possible true positive	58-470
2D	Mus musculus dynein, cytoplasmic, light chain 1	Correct	Coding	1	Yes	Possible true positive	28-470

Table 3.3.1.7 Results of Yeast Two Hybrid screen using the N terminus mTbx1 as bait.

3.4 Discussion

Tbx1 is a transcription factor and it is therefore likely to have protein binding partners that act as cofactors in transcription. Transcription factor proteins have protein binding partners that are important for their transcriptional activity. It is known that Tbx1 acts in synergy with Nkx2.6 to activate the ANF promotor (Heathcote et al. 2005). Tbx1 has also recently been found to act in synergy with Pitx2 (Nowotschin et al. 2006).

As the yeast two hybrid system is an artificial representation of the *in vivo* situation, positive clones should be examined carefully for their plausibility as protein binding partners for Tbx1. Factors such as spatial and temporal protein expression patterns, protein function and cell localisation should be considered when examining clones as potential candidates for interaction. Also protein credibility can be assessed by searching the literature for known binding partners of candidate proteins.

3.4.1 Potential protein-protein binding partners of the GAL4-mTbx1 C terminus fusion protein

The mTbx1 C terminal bait fusion protein was detected by Western analysis using an anti-GAL4 antibody. This therefore suggested that the construct was expressed in the yeast two hybrid system. The large number of colonies obtained (in excess of 1500 with one tenth the amount of bait protein) suggested that the C terminus of mouse Tbx1 has many potential binding partners. This high colony number could also represent a high number of false positives causing expression of the reporter genes. As the C terminus contains an activation domain it would be predicted to interact with more proteins than the N terminus. This notion is supported by the small colony numbers obtained for the mTbx1 N terminal protein.

Table 3.3.1.6 shows the list of VP16 library clones that survived on the selection plates and were successfully sequenced from the C terminus library screen. Proteins were excluded as potential interactors if they were from the noncoding region of the gene, the incorrect orientation for expression, or if they were out of frame with GAL4. This excluded all but three clones: Jag1, Kielin, and Granulin. Kielin is a secreted signalling

molecule that mediates inductive activities of the embryonic midline. Secreted proteins are poor candidates for interactions with transcription factors, which are presumed to have mainly nuclear expression. Thus Kielin and Granulin (also a secreted protein) are likely false positives. Jag1 is a ligand that is mutated in Alagille syndrome (Li et al. 1997a). Patients with this syndrome have right sided heart defects, which make Jag1 a good candidate for a Tbx1 protein interactor. Jag1 does have a transmembrane domain and a cytoplasmic tail, but the region of interaction detected by the yeast two hybrid screen was extracellular. These three candidates were therefore not considered strong enough for further work.

Although 357 colonies were present after the second round of selection for the C terminal construct, only 30/357 (8%) of these were successfully identified by PCR and sequencing (table 3.3.1.5). This compares to 16/18 (89%) plasmid isolation from the N terminus construct. There are several possible reasons for this. The first is autoactivation by the C terminal construct. This could account for the high number of positive colonies (1500) with just one tenth of the library screen. Also this might be anticipated as the C terminus of Tbx1 is thought to confer transcriptional activation. On re-inspection of the autoactivation plates (figure 3.3.1.3b) there does appear to be more growth on the SD-uracil-adenine plates for the C terminus compared to the N terminus, but at the time this was not considered to be enough to represent autoactivation with cessation of the experiment.

3.4.2 Potential protein-protein binding partners of the GAL4-mTbx1 N terminus fusion protein

The N terminal yeast two hybrid screen was conducted at ten fold the scale of the C terminal screen. Despite this, the number of colonies remaining after histidine and adenine selection was only 18 (table 3.3.1.5). It is noteworthy that the N terminus fusion protein was not detected on Western blot analysis possibly indicating a low level or transient expression of the protein. This might suggest that positive clones from the screen are stronger interactors as there is less bait protein to bind to. 18 clones grew in the absence of histidine and adenine indicating an interaction between the N terminal mTbx1 fusion construct and a VP16 fusion protein. All except one of these clones carried prey plasmids encoding the open reading frame of Dynein light chain 1, cytoplasmic, also named PIN (protein inhibitor of neuronal nitric oxide synthase), LC8,

DLC1 and DNCL1. DLC1 is one of the most evolutionarily conserved sequences yet described. It is identical at least in mouse, rat, pig, cow and human and is ubiquitously expressed. The identification of four independent clones reduces the chances of the interaction being a false positive.

DLC1 has been mainly implicated in several cytoplasmic transport processes (Wilson et al. 2001). It is an essential component of the cytoplasmic dynein complex and associates with other cytoplasmic molecules. To date more than 10 different molecules of diverging sizes and biological functions have been found to interact with DLC1. Most of these proteins have no apparent link with each other. Some of these protein interactions seem to be associated with minus-end directed microtubule-based intracellular transport. In parathyroid cells DLC1 interacts with the 3'-untranslated region of the parathyroid hormone mRNA to translocate it along microtubules for its further processing or translation (Epstein et al. 2000a).

Other DLC1 protein interactors include IκBα an inhibitor of the NFκB transcription factor by promoting its intracellular retention (Crepieux et al. 1997). DLC1 is also known to associate with chicken brain myosin-Va, an actin based molecular motor (Espindola et al. 2000).

DLC1 has more functions besides being a simple link molecule in intracellular transport. It has been shown to interact with BIM (bcl-2 interacting protein) and is able to regulate its pro-apoptotic activity (Yu et al. 2002). There is also evidence for a nuclear function of DLC1 and its association with various transcription factors. Herzig *et al* used the N terminus of the transcription factor nuclear respiratory factor-1 (NRF-1) as bait and screened an E9.5/E10.5 embryonic mouse VP16 library and a mouse testes library and obtained DLC1 as a positive interactor from both libraries (Herzig et al. 2000). Fluorescence overlays of confocal images established that NRF-1 and DLC1 display a very similar nuclear staining pattern. The interaction was also detected by co-immunoprecipitation reactions. In addition the group used the drosophila erect wing gene (EWG) product, a protein related to NRF-1 through its DNA binding domain, to show that this too interacted with DLC1 *in vitro* and in transfected cells.

Kaiser *et al* used the C terminus of TRPS1 as bait in a yeast two hybrid screen (Kaiser et al. 2003). *TRPS1* is a gene mutated in tricho-rhino-phalangeal syndrome type 1

(Momeni et al. 2000). Two independent clones containing DLC1 were obtained from the screen and the interaction domain was found to be within the zinc finger motifs and the IKAROS-like motif of TRPS1. Evidence for the interaction was strengthened using an *in vitro* glutathione-S-transferase (GST) pulldown assay. The interaction was confirmed in primary human fibroblasts by co-immunoprecipitation, co-localisation, DNA binding and functional assays. Importantly DLC1 was shown by the group to be expressed in the nucleus. A GATA regulated reporter gene assay indicated that DLC1 suppresses the transcriptional repression activity of *TRPS1* (Kaiser et al. 2003).

In another yeast two hybrid system Jaffrey and Snyder used human neuronal nitric oxide synthase (NOS) as bait and isolated a cDNA interacting clone containing PIN (Jaffrey & Snyder 1996). They showed that PIN physically interacts with and inhibits the activity of neuronal NOS. PIN is the same protein as DLC1 and may be involved in apoptosis, synaptogenesis and neuronal development. The results suggest that PIN may regulate numerous biological processes through its effects on nitric oxide synthase activity.

More recent work has provided further support for the association of DLC1 with nuclear proteins. Cooper *et al* have shown that DLC1 interacts with PAX6 (Cooper & Hanson 2005) a homeodomain transcription factor. In addition Lo *et al* have shown that DLC1 binds p53 binding protein 1 (53BP1) and mediates DNA damage induced p53 nuclear accumulation (Lo et al. 2005). Rayala *et al* suggest a potential chaperone-like activity of DLC1 in the nuclear translocation of the oestrogen receptor (ER). This follows their observation that DLC1 interacts with the ER and its downregulation led to compromised ER transactivation activity and its nuclear accumulation (Rayala et al. 2005).

The known binding domains of DLC1 targets are confined to short stretches of amino acid residues. For instance a GIVQD amino acid sequence was recognised in the DLC1 interacting region of nNOS and a GKAP stretch in BIM. The group that identified DLC1's interaction with 53BP1 isolated two binding domains: GIQTM and AATQT (Lo et al. 2005). No previously described interaction domains are present in the N terminus of mouse Tbx1.

As DLC1 is expressed in the nucleus and is known to interact with and suppress other transcription factors, this protein is a possible candidate for interactions with Tbx1. It has also been previously reported that Tbx1 can localise to the cytoplasm (Goding 2002). As DLC1 has a diverse range of functions both in the nucleus and the cytoplasm, it is plausible that it interacts with Tbx1 and modifies its transcriptional activity.

3.4.3 Future Work

The yeast two hybrid screen has provided a possible candidate in DLC1 for an interaction with the mouse Tbx1 N terminus protein. Clearly further work is needed to verify this interaction (discussed in Chapter 8). This work was performed concurrently with Affymetrix microarray studies (Chapters 4-6). Due to time constraints, efforts were focussed on the latter avenue of study as results were more encouraging.

Chapter 4 Identification of *Dfl* target genes

4.1 Introduction

The transcription of genomic DNA into mRNA is the first step in the process of protein synthesis, and differences in gene expression in different biological settings are useful in the understanding of molecular processes. High density synthetic oligonucleotide arrays provide a tool to analyse the expression of thousands of genes expressed in one tissue at one time. By comparing expression levels in a control situation with levels in an altered biological situation, differentially expressed genes can be examined more closely for their role in the altered process.

4.1.1 Microarrays in general

There are many different types of arrays available for both DNA and mRNA analysis. It is now possible to perform a genome wide linkage analysis screen with the use of a DNA SNP (single nucleotide polymorphism) microarray and mutation analysis can be performed more quickly by ordering region specific oligo (oligonucleotide) arrays.

Various methods are available for detecting and quantitating gene expression levels including Northern blots (Alwine et al. 1977), representational difference analysis (Hubank & Schatz 1994), subtraction hybridisation and serial analysis of gene expression (Velculescu et al. 1995; Velculescu et al. 1995). Augmenting these methods are two array based technologies: cDNA and oligo arrays.

Spotted cDNA arrays necessitate fluorescent labelling of control and sample mRNA with Cye3 or Cye5, co-hybridisation of samples and scanning at two different wavelengths to detect relative transcript abundance for each condition. Arrays can be purchased with a cDNA population representing most of the genome or they can be designed to look at a specific disease process such as cardiac development for example. There can however be problems with clone contamination and cross hybridisation, and the set up is often slow and expensive.

Oligo arrays are more usually employed to examine expression analysis. Oligonucleotide arrays provide good chip to chip reproducibility, but are less flexible

than the cDNA spotted arrays. They also have the advantage that less total RNA is needed to start with and that temporal gene expression changes can be analysed.

4.1.2 The Affymetrix Genechip System

4.1.2.1 Microarray manufacture and probe design

Affymetrix Genechip™ probe arrays are manufactured by using photolithography where oligos representing 12,000 genes are synthesised on a glass base. This technique allows approximately 0.1pmol of oligos to be synthesised per mm^2 of glass. Each gene is represented by 16 probe pairs consisting of 25mer oligos positioned on different areas of the arrays. Each probe pair consists of a perfect match oligo and a mismatch oligo with single base change positioned at the 13^{th} base pair. This mismatch probes act as specificity controls that allow direct subtraction of both background and cross-hybridisation signals and allows determination of non-specific binding levels (www.Affymetrix.com).

4.1.2.2 Sample preparation for Genechip arrays

Figure 4.1.2 summarises the steps involved in sample preparation for GeneChip arrays. Firstly total RNA is extracted from the sample tissues, cleaned, precipitated and quantified. Secondly the total RNA is converted to single stranded cDNA using reverse transcriptase and a poly-T primer, and then to double-stranded cDNA using DNA polymerase 1 and DNA ligase. The sample is cleaned and precipitated ready for the *in vitro* transcription (IVT) labelling reaction. The IVT reaction involves amplification of the cDNA using a T7 RNA polymerase in the presence of biotin-UTP and biotin-CTP, so each cDNA will yield 50-100 copies of biotin labelled cRNA. This cRNA is subsequently fragmented to yield fragments of between 35 and 200 nucleotides in length. The cRNA fragments are hybridised to the microarray chip and then stained with Streptavidin-Phycoerythrin and washed. The signal on the chip is amplified with goat IgG and biotinylated antibody, and the intensity level of this signal is recorded as the chip is scanned by a confocal laser scanner (www.Affymetrix.com).

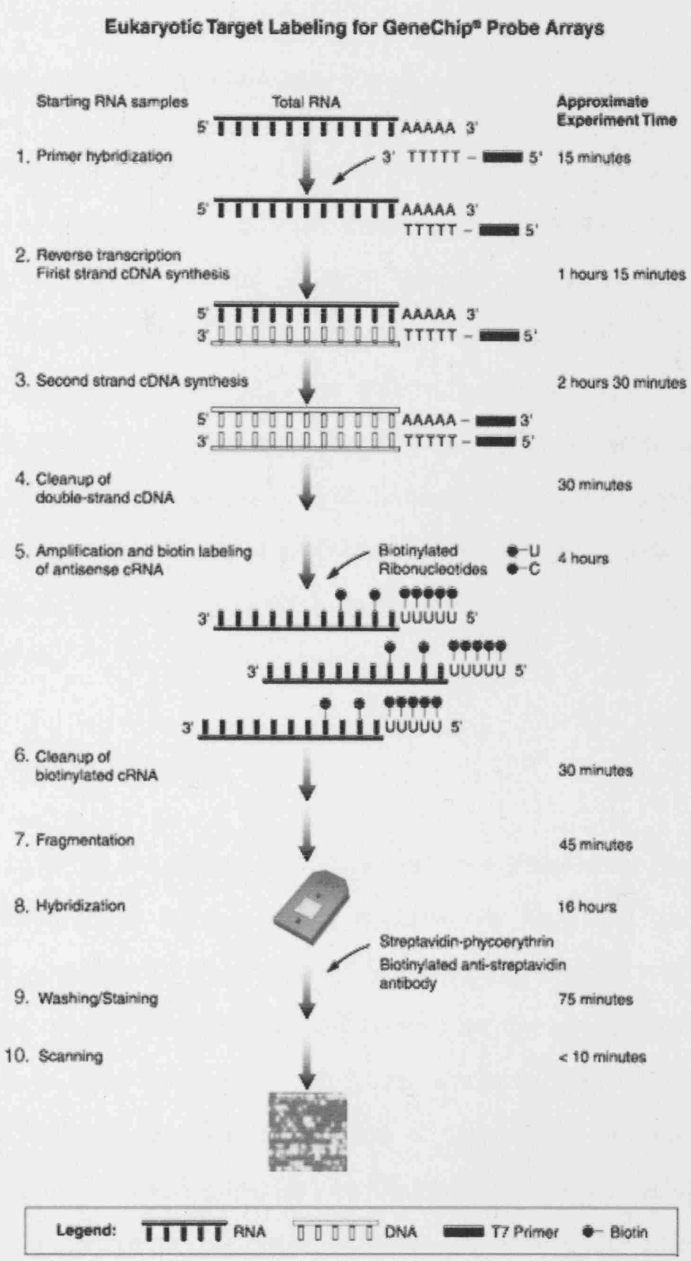


Figure 4.1.2.
Illustration of the methodology for sample preparation for Affymetrix Genechip hybridisation. Briefly, total RNA is reverse transcribed using a T7-oligo(dT) primer, and double-stranded cDNA is synthesized. The double-stranded cDNA with incorporated T7 promoter is then used as a template in the subsequent *in vitro* transcription reaction. Figure from the Affymetrix website: www.Affymetrix.com.

4.1.2.3 Quality control employed by Genechip microarrays

A quality control technique employed by Affymetrix is the use of spiked controls within the cRNA mixture. The hybridisation mixture consists of non-eukaryotic hybridisation controls. These control cRNAs are spiked in at specific concentrations to allow an interpretation of the labelling, hybridisation and scanning of each array to ensure array experiments are comparable. "Absent calls" or relatively low signal values from these controls indicate a potential problem with the hybridisation reaction or subsequent washing or straining steps. The eukaryotic hybridisation controls are bioB, bioC, bioD from biotin synthesis of *E.Coli* and *cre* from P1 bacteriophage. A synthetic biotinylated control oligo (Oligo B2) is also added to the hybridisation solution to provide alignment signals for image analysis.

4.1.3 Application of microarrays

Microarrays can be used to examine differential gene expression in many biological conditions. For instance some of the earlier applications involved looking at gene expression in yeast under different environmental conditions. Pharmaceutical industries have employed microarrays to identify potential targets for therapeutic intervention (Shi 2001). Microarrays have been particularly useful to cancer biologists both in studying oncogenic pathways and establishing blueprints of particular carcinogenic tissues. Results from these studies can aid clinicians in diagnosis of cancer subtypes to allow fine-tuning of chemotherapy treatment options (Olson, Jr. 2004). Use of microarrays to analyse mouse embryonic tissues is still in its infancy, but it is a rapidly expanding field with great potential in the analysis of biological processes.

4.1.3.1 The use of microarrays in mammalian development

There is an abundance of literature examining expression profiles with microarrays in *Drosophila* and *C.Elegans*, but a relative paucity in mammalian systems. The majority of the latter utilise transfected or treated cell lines such as *Pax3* (Mayanil et al. 2001) and *Tbx2* (Chen et al. 2001) to identify potential targets. Exploitation of mouse mutants in microarray analysis is also in its infancy, with the vast majority using differential expression in *homozygous* rather than heterozygous mutants compared to wild types (Kennan et al. 2002; Hedlund et al. 2004). Many mutant homozygous embryos have

such a severe phenotype that certain tissues present in the wild type embryos are lacking in the mutant embryos. This can add the additional complexity of false positive microarray results as differentially expressed genes may be due to absent or deranged tissue in the mutant embryo and not represent true targets. This can often be overcome by using heterozygous mutant samples that have less structural defects, although the corollary of this is that heterozygous gene expression changes will inevitably be more subtle.

4.1.3.2 Application of Genechips for differential gene expression in Df1 embryos

To test the reliability of gene expression changes in a heterozygous system, the *Df1* mouse model has been used to analyse differential gene expression at the mouse E10.5 stage. Use of the *Df1* mouse model has a distinct advantage over most other mouse mutants in that there are 22 genes within the deleted region that can be exploited to assess the reliability of expression data. These genes were used as "in built" positive controls when performing filtering and statistical analysis of data.

The Ts65Dn mouse is an example of a copy number mutant with a change in its transcriptome analysed by microarray. This Ts65Dn mouse parallels human trisomy 21 (Saran et al. 2003). In this model for Down syndrome there is a dosage imbalance of 124 genes on the distal end of mouse chromosome 16 which shows near perfect conserved synteny with human chromosome 21. Interestingly microarray analysis of this model revealed subtle changes in individual genes, but cluster analysis revealed a more general disturbance of 7000 genes located throughout the genome (Saran et al. 2003).

4.1.3.3 Aims of target identification in Df1 embryos

In the light of the advantages and limitations of the *Df1* model, the aims of this experiment were several fold. The primary aim was to identify significant changes in gene expression that might represent a primary or secondary response to the *Df1* deletion: so called "target identification". In addition, I asked whether hemizygosity for *Df1* genes resulted in a significantly greater than two-fold change in the expression level of any other gene represented on the array. This phenomenon is termed "amplification of haploinsufficiency", and it has been observed previously, most notably for *Tbx5*

(Bruneau et al. 2001). Secondly, I wished to determine whether the *Df1* deletion affected the expression of genes adjacent to, but not within the deletion. The presence of non-overlapping rearrangements of 22q11 in patients with 22q11DS has led to the proposal that some of these chromosome abnormalities act via a long range effect on transcription and that the aetiology of 22q11DS is complex (Botta et al. 1997; Amati et al. 1999; Novelli et al. 2000). This has been observed for *Pax6* in aniridia and *Sox9* in campomelic dysplasia, autosomal sex reversal (Kleinjan et al. 2001; Pfeifer et al. 1999). Of the genes in proximity of the deletion, *Hira* is of particular interest since knockdowns in a chick system have been associated with defects in outflow tract septation, and null mutants die prior to the development of structures affected in 22q11DS (Farrell et al. 1999).

Finally, in cases where genes deleted in *Df1* were not apparently downregulated on the microarray, my aim was to investigate whether this could be explained by dosage compensation or upregulation of the non-deleted allele. Dosage compensation could partly explain the lack of haploinsufficient phenotypes for the majority of genes in the genome, and the analysis of *Df1* deleted genes gave the opportunity to assess how widespread this mechanism might be.

4.2 Experimental methodology

Figure 4.2.1 summarises the methodology employed for analysis of the gene expression profiles in the *Df1* mouse. The *Df1* mouse carries a genetically engineered 1.2 Mb deletion on the syntenic chromosome region to human 22q11.2. As the majority of the pathology in *Df1* mice involves the branchial arch branchial arch dissections of the 1st-6th arches in wild type and *Df1* mice were used to examine differential gene expression. Embryonic day 10.5 embryos were used as at this stage there is 100% penetrance of the hypoplastic 4th arch artery defect which is rescued in later embryos.

To minimise experimental error, and to gain enough RNA for analysis, I opted to pool three embryos at the E10.5 stage. In addition due to the anticipated subtly of the changes, I opted to perform analysis using 7 different embryo triplicates each of wild type and *Df1* to improve statistical reliability of the data. To reduce sample variability the embryos were dissected at the same time of the day using the same technique and sibling pairs were used in the experiments wherever possible.

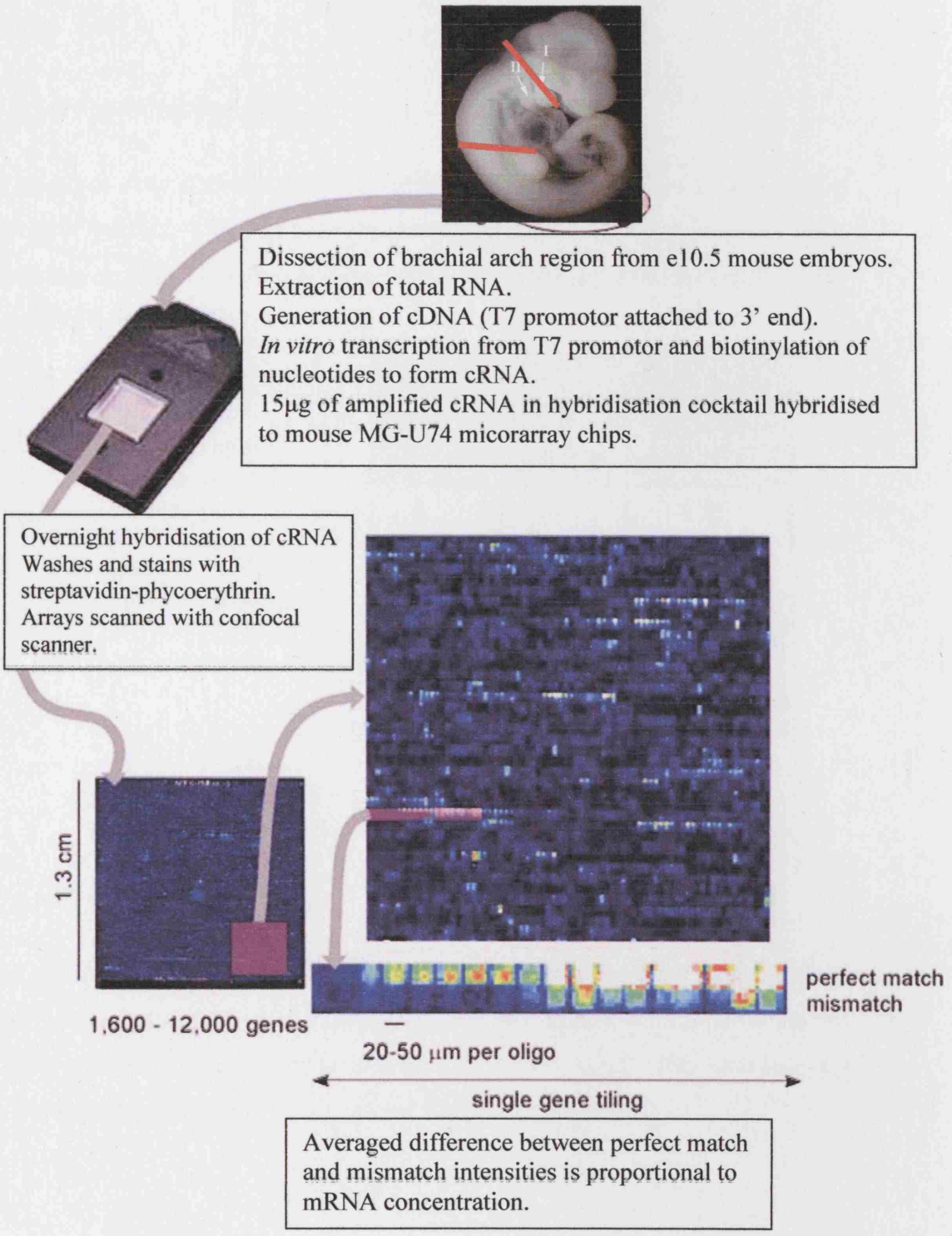


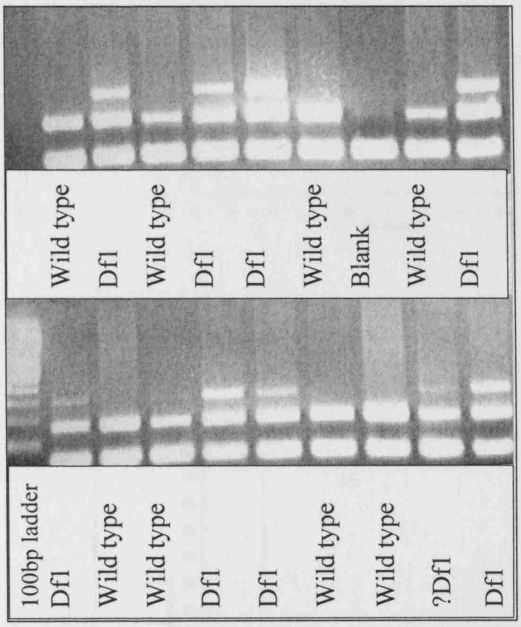
Figure 4.2.1
Overview of gene expression profiling using Affymetrix arrays. Hybridisation intensities are shown as false-colour scale from blue (low) to white (high). Genes scored as present show higher intensity hybridisation to the perfect match oligos than to the mismatch oligos. Figure adapted from www.Affymetrix.com.

4.3 Results

4.3.1 Embryo genotyping PCR

Once the embryos were dissected the yolk sac was used to extract DNA for genotyping to differentiate between wild type and *Df1* deleted samples. An example of the three primer genotyping PCR is shown in figure 4.3.1 and embryos with equivocal results were excluded from the study.

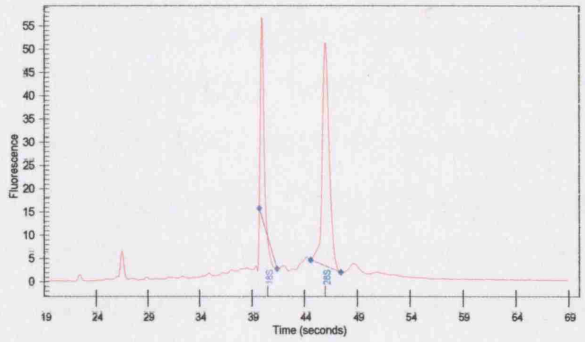
Figure 4.3.1
Example of 3 primer PCR used to genotype the *Df1* embryos.
Df1: *Df1*/+
Wild type: +/+
?*Df1*: equivocal result

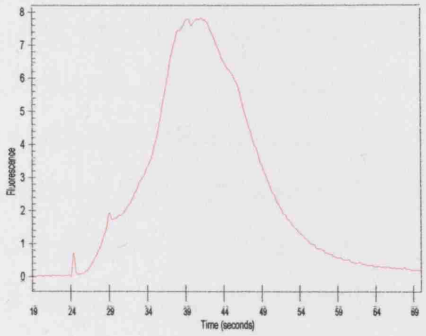


4.3.2 Quality and yield of RNA

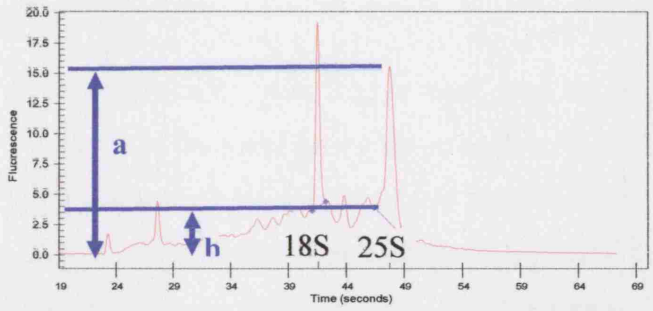
The RNA extracted from the branchial arch dissections was analysed on the Agilent bioanalyser to assess its quality before and after labelling. Figure 4.3.2 shows examples of total RNA and cRNA profiles in samples of high and low quality that were assessed for array hybridisation. Samples with degraded RNA were not used.

Sample 1 **Total RNA** **Post IVT**





Sample 2 **Total RNA** **Post IVT**



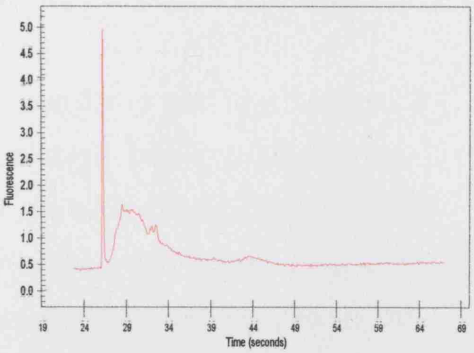


Figure 4.3.2.
Examples of two RNA samples assessed on the Agilent bioanalyser. The graphs on the left denote the samples as total RNA and the graphs on the right show the samples after the *in vitro* transcription (IVT) reaction has produced cRNA. The X axis represents time in seconds and the Y axis represents fluorescence detected by the Agilent machine. Sample 1 illustrates an example of good quality total RNA and cRNA. However sample 2 is an example of rejected RNA because there is some evidence of degradation in the total RNA. Also the cRNA post IVT shows low abundance with short products also illustrating degradation. The labels on the sample 2 total RNA show an approximate method of assessing degradation. It is generally accepted that if a/b is >8 then the degradation is minimal and the sample can be used. In sample 1 a/b=12.5; in sample 2 a/b=4.

4.3.3 MG-U74AvMicroarray Hybridisation

Wherever possible microarray chips from the same batch were used for comparable experiments and chips with poor hybridisation parameters were rejected. Table 4.3.4 depicts a summary of the report files which represent the hybridisation parameters for each chip. A target intensity of 100 was arbitrarily chosen for each chip. The scaling factor (SF) is a measure of the amount the signal intensities had to be adjusted for each chip to provide a target intensity of 100. The Affymetrix manual suggests the background (BG) should be less than 100 which is a reflection of the brightness of the chip when scanned. Chips with a background intensity greater than 100 were excluded from the data. The Raw Q is influenced by the electrical noise of the scanner and the sample quality. Chips with a Raw Q greater than 4 in the report files were also excluded. Ideally the present/absent calls of the genes should be in a ratio of 50%:50% and sometimes a low present call can reflect a high scaling factor. The 3'/5' ratio represents the relative abundance of RNA transcripts detected and should ideally be around 1. Higher figures show possible sample degradation.

Ten good quality microarrays were used for analysis containing pooled RNA from 3 embryos of the same genotype on each (n=5 wild type embryos and n=5 *Dfl* embryos). In addition four microarrays were hybridised by Dr Sarah Ivins (MMU) using labelled cRNA from Elisabeth Lindsay, Baylor, Houston that was isolated in the same way from E10.5 wild type (n=2) and *Dfl* (n=2) embryos. The results from both the experiments were combined to improve statistical power.

4.3.4 Rejected and borderline microarray chips

In total, two wild type microarray chips were rejected due to a background reading greater than 100. The details of these chips are shown in table 4.3.4 below. The chip with the background reading of 1900 showed up as a complete white-out when scanned. It is unclear why chips have a high background. One theory is that there is ethanol contamination of the sample but it could also be a fault with the arrays themselves. As can be seen from table 4.3.4, having a high background tends to reduce the present call percentage.

Sample Df1D1 had a scaling factor of 9.1 which is around the borderline figure for rejection (the Affymetrix manual states that scaling factors should be within three-fold of each other). This suggests there was a reduced amount of RNA hybridised to the chip. Sample Df1D6 has a relatively high 3'/5' GAPDH figure of 3.82 suggesting some RNA degradation. In addition sample Df1Wt6 has a relatively low present call of 30.1% suggesting low signal intensities perhaps reflecting poor probe labelling. I decided that these chips should be included in the analysis as the results from the *Df1* control genes were within similar limits to other *Df1* chips when compared to wild type samples.

Chips with acceptable parameters and therefore included in data analysis:

	RNA Type	SF	Raw Q	BG	%P	%A	3'/5'(GAPDH)
Df1D1	3 *Df1* embryos	9.10	1.57	34	31.2	66.3	1.12
Df1D2.2	3 *Df1* embryos	1.43	2.43	69	47.4	50.6	0.96
Df1D4	3 *Df1* embryos	3.18	2.46	62	36.4	61.4	1.79
Df1D5	3 *Df1* embryos	2.90	2.20	52	39.0	58.0	1.99
Df1D6	3 *Df1* embryos	3.79	2.00	49	37.3	60.4	3.82
MA3	3 *Df1* embryos	1.02	2.02	50	51.5	46.5	2.49
MA4	3 *Df1* embryos	1.09	2.50	64	49.1	49.0	2.41
Df1Wt1	3 Wt embryos	2.50	1.40	32	48.4	49.5	0.98
Df1Wt2	3 Wt embryos	3.70	1.50	35	43.4	54.4	1.11
Df1Wt3	3 Wt embryos	3.30	1.60	35	44.0	54.0	1.10
Df1Wt5	3 Wt embryos	4.99	2.30	58	37.5	55.5	1.07
Df1Wt6	3 Wt embryos	5.66	2.21	53	30.1	67.3	2.80
MA5	3 Wt embryos	1.23	1.79	41	52.9	45.2	1.12
MA6	3 Wt embryos	1.02	1.78	42	51.7	46.2	1.10

Chips rejected from data analysis due to unacceptable quality parameters:

	RNA Type	SF	Raw Q	BG	%P	%A	3'/5'(GAPDH)
Df1Df2.1	3 *Df1* embryos	1.62	4.7	161	30.0	70.0	1.08
Df1Wt4	3 Wt embryos	0.90	22.0	1900	3.4	96.2	1.00

Table 4.3.4

Summary of the report files of the microarray chips used in the data analysis. MA3-MA6 represent the chips hybridised by Dr Ivins. **SF**: scale factor, **BG:** background

intensity, **%P**: % genes called as present, **%A**: % genes called as absent, **3'/5' GAPDH**: ratio of 3' and 5' transcripts of GAPDH used as a measure of degradation. Shaded samples were either rejected or borderline for rejection (see text).

4.3.5 Results of Genespring analysis

4.3.5.1 Elimination of genes through filtering and statistical analysis

Microarray data is too complex to understand in its raw form, necessitating a filtering and statistical approach to produce a meaningful gene list for potential follow-up. The data handling and filtering employed in Genespring is modified from the recommended approach for analysing replicate chips from two different conditions: in this case wild type and *Df1* embryonic RNA. Figure 4.3.5.1 illustrates the process of data handling in Genespring.

The 14 metric text files (seven wild type and seven *Df1* files) from Affymetrix analysis were imported into Genespring version 4.2.1 (Silicon Genetics, Inc, Redwood City, California). Wild type signal intensities were normalised to one to allow a more simplistic approach to data analysis. The data from the seven *Df1* chips were averaged for the data interpretation (Fig 4.3.5.1 data interpretation method A). Filtering and statistical analysis was applied to the data and the numbers of genes eliminated at each stage is shown in figure 4.3.5.2. The total number of genes represented on the mouse MG-U74Av microarray chip is 12,488. Approximately a third of these were rejected as they had control strengths below the assigned lower limit of acceptance for this dataset in Genespring (figure 4.3.5.2). This eliminates transcripts that have low hybridisation levels and are therefore untrustworthy for statistical expression analysis.

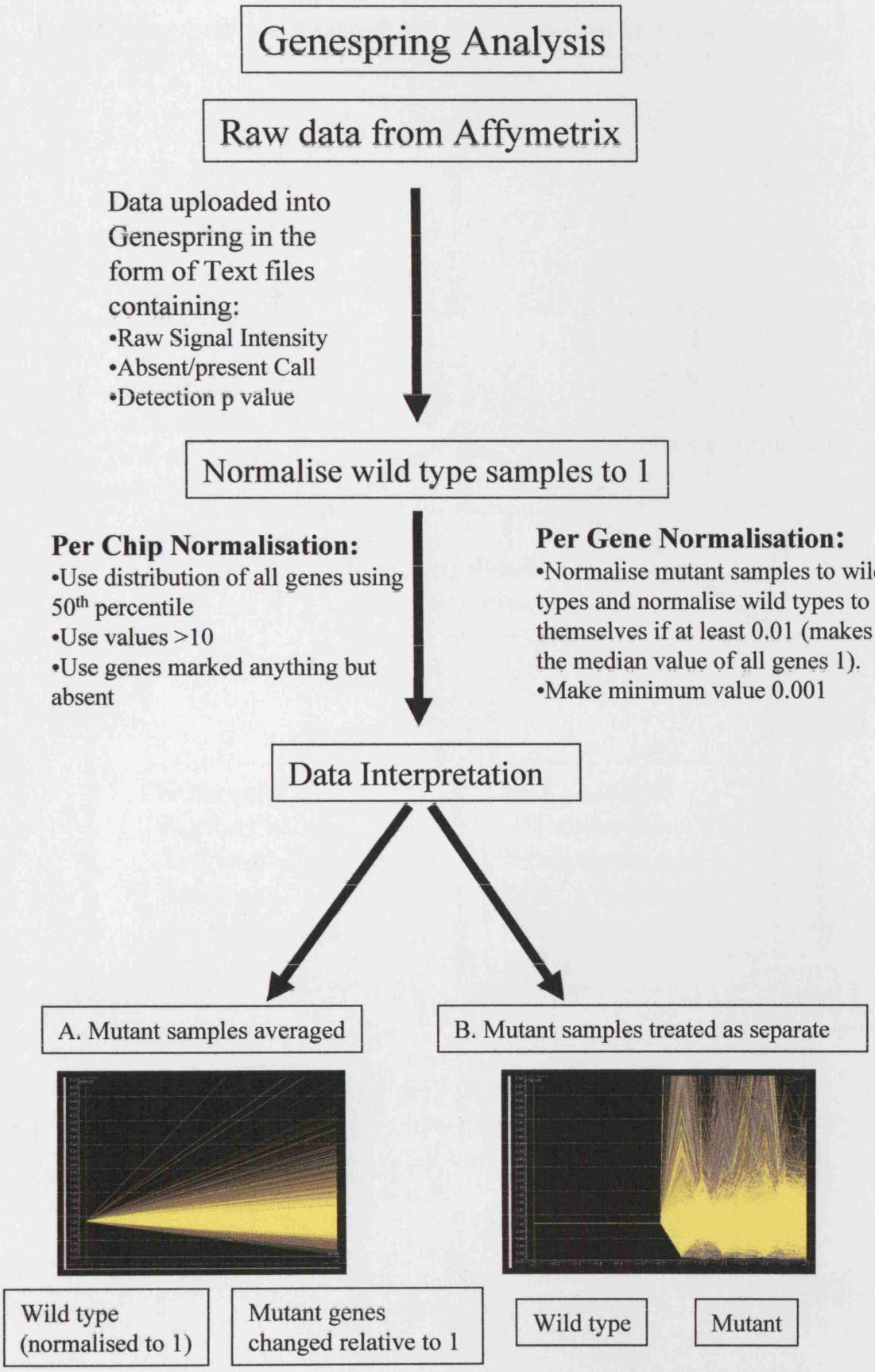


Figure 4.3.5.1.
Diagram illustrating the process of data handling in Genespring. Graphs illustrate two possible methods of data interpretation where the yellow lines represent genes, the Y axis is the normalised intensity ratio and the X axis represents the wild type and mutant datasets.

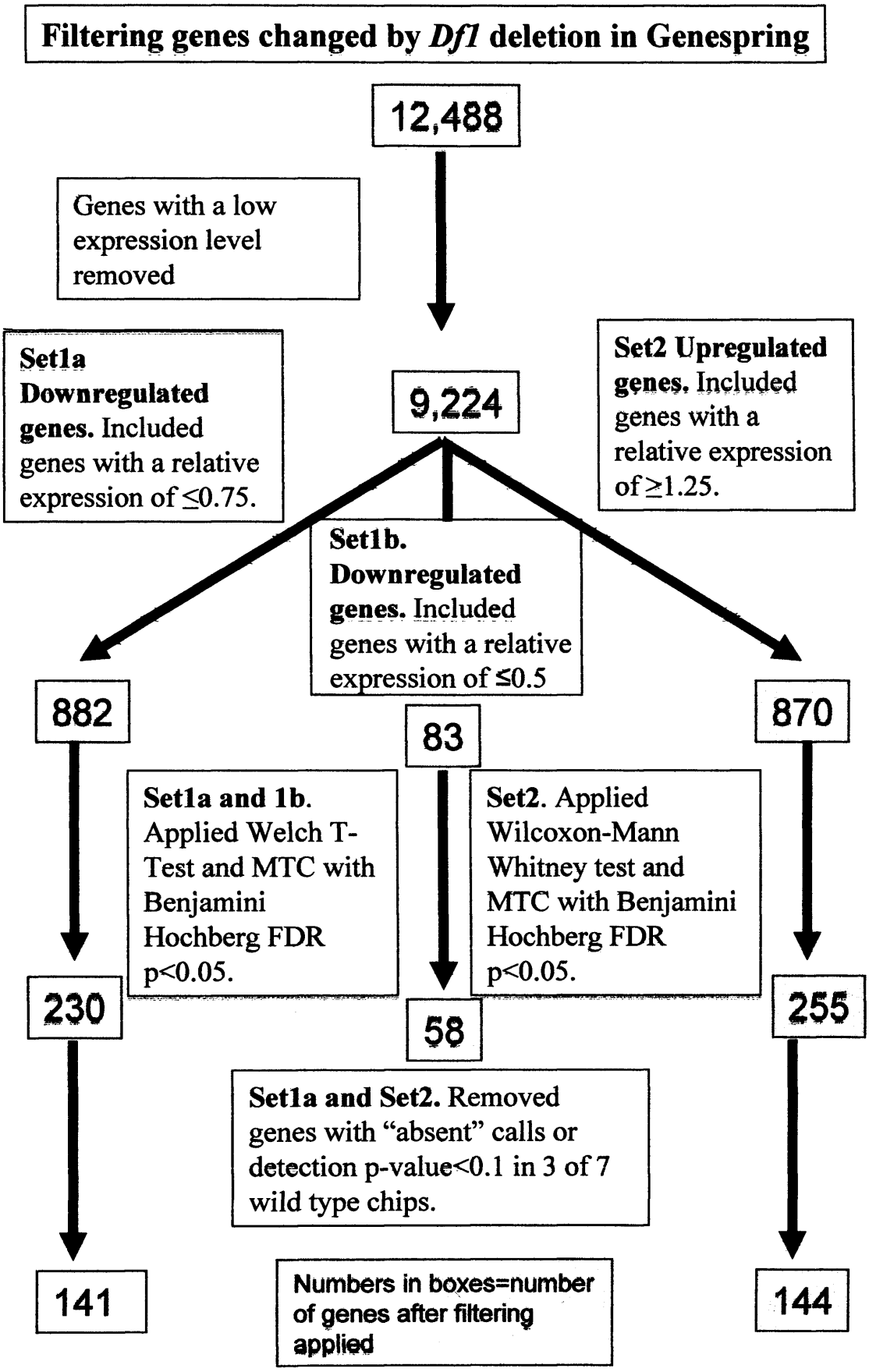


Figure 4.3.5.2
Diagram representing the number of genes eliminated at each stage of filtering in Genespring in the Df1 dataset where the wild type chip data is normalised to 1. MTC: multiple testing correction; FDR: false discovery rate.

4.3.5.1.1 Filtering the downregulated gene sets: 1a and 1b

The next filtering constraint applied was expression level. For the downregulated genes (Set 1a) 0.75 was used as the cut-off value; therefore genes in the *Df1* dataset that were 75% of wild type levels and below were included. This left 882 genes fulfilling the above combined criteria. *Df1* embryos are heterozygotes and therefore the genes within the deleted region are present at only 50% at the genomic level. It would therefore be expected that these genes would come through the data with approximately 50% expression levels compared to the wild type controls. The 0.75 cut off value was chosen to allow inclusion of genes downregulated by the deleted genes by 25% or more and it was considered that this degree of downregulation might represent biological significance.

Figure 4.3.5.2 also shows the results of applying the Welch t-test (parametric, variances not assumed to be equal) and multiple testing correction (MTC) using Benjamini Hochberg (BH) false discovery rate (FDR) at p values less than 0.05. This resulted in a downregulated list of 230 genes. Finally only those genes with a detection p value less than 0.1 in the wild type data were included resulted in 141 genes. This last filter excludes genes that have low perfect match/mismatch ratios which could give rise to misleading results.

A second set of changed genes (set 1b) was obtained with a view to highlighting potentially relevant genes of known biological function with larger fold changes, but missed in the first selection through application of the detection p value filter in Genespring. This set was derived by eliminating genes with a low control strength, applying a relative downregulation cut-off of 0.5 and filtering using the Welch t-test (variances not assumed to be equal) using the BH FDR multiple testing correction ($p<0.05$). The detection p value filter was omitted as it eliminated most of these genes. In total 83 genes were detected with expression levels less than 0.5. 58 genes passed through the Welch T-test and MTC with BH FDR $p<0.05$. Table 4.3.5.1.1 shows downregulated gene set 1a including some genes from set 1b.

Table 4.3.5.1.1

Affymetrix No.	Genes implicated in vasculogenesis/angiogenesis	Df1	p value
96679	DnaJ (Hsp40) homolog, subfamily B, member 9	0.46	0.0275
104385	integrin alpha V	0.58	0.0304
104590	myocyte enhancer factor 2C	0.58	0.0398
100947	transcription factor 20	0.62	0.0326
104635	gap junction membrane channel protein alpha 7-connexin 45	0.65	0.0174

Affymetrix No.	Genes implicated in cardiogenesis	Df1	p value
102419	retinoic acid receptor, gamma	0.55	0.0128
93667	F-box and WD-40 domain protein 7, archipelago homolog (Drosophila)	0.59	0.0275
93669	SRY-box containing gene 11	0.68	0.0442

Affymetrix No.	Transcription factors/cofactors	Df1	p value
102085	insulinoma-associated 1	0.286	0.0111
96500	Gata5*	0.36	0.018
93950	forkhead box A2 (Foxa2)*	0.37	0.004
96169	sal-like 4 (Drosophila)	0.47	0.0304
92747	NK2 transcription factor related, locus 2 (Drosophila)	0.58	0.0442
93701	Smarca5	0.59	0.0048
103302	SRY-box containing gene 3	0.6	0.0243
161146	T-box 2	0.62	0.0365
101684	M.musculus T2 mRNA.	0.66	0.025
103288	nuclear receptor interacting protein 1	0.68	0.0119
102069	metal response element binding transcription factor 2	0.68	0.0415
160498	LIM domain binding 1	0.7	0.0275
103301	SRY-box containing gene 3	0.72	0.032
104645	Kruppel-like factor 7 (ubiquitous)	0.72	0.032
94176	LIM homeobox protein 6	0.72	0.0326
92912	homeobox protein; M.musculus Msx-3 gene, exon 2.	0.72	0.0448
102303	zinc finger protein 61	0.73	0.0279

Affymetrix No.	Protein Signalling pathway genes	Df1	p value
97227	guanine nucleotide binding protein, alpha 12	0.35	0.0415
101831	Sonic hedgehog (Shh)*	0.45	0.012
93733	regulator of G-protein signaling 19 interacting protein 1, synectin	0.5	0.0355
98602	RAN GTPase activating protein 1	0.59	0.0151
98001	Rho guanine nucleotide exchange factor (GEF) 1-signal transduction	0.61	0.0047
94262	thyroid hormone receptor associated protein 3	0.62	0.033
99975	protein-kinase, interferon-inducible double stranded RNA dependent inhibitor	0.63	0.0063
104506	nuclear receptor subfamily 1, group I, member 3	0.67	0.0103
161161	expressed in non-metastatic cells 1, protein	0.67	0.0304
97226	guanine nucleotide binding protein, alpha 12	0.71	0.0145
98084	ADP-ribosylation-like 2 binding protein	0.72	0.0146
94491	peroxisomal biogenesis factor 13	0.73	0.0053

Table 4.3.5.1.1 Downregulated genes. Set1a.

Genes downregulated in the Df1 compared to the wild type cDNA samples and passing the filtering and statistical analysis described in the methods as gene set1a. Genes with * are from set1b (see methods). ESTs and genes expressed at 0.71-0.75 wild type levels are not shown in the table to conserve space. **Affy no**: affymetrix number, **MA**: relative expression in the Df1:wild type microarray chips averaged across the 7 arrays. nb Df1 genes are shown in a separate table.

Affymetrix No.	Cytoskeletal genes	Dfl	p value
95466	coactosin-like 1 (Dictyostelium)	0.42	0.0048
100923	myosin X	0.6	0.0151
96033	M.musculus syndecan-1.	0.71	0.0313
92346	Mouse NF-M gene for middle-molecular-mass neurofilament protein.	0.65	0.0352
103330	spermatid perinuclear RNA binding protein	0.68	0.0191
102108	myosin heavy chain IX	0.71	0.0475
94270	Mus musculus cytokeratin (endoB) gene, complete cds.	0.72	0.0174
95637	filamin, beta	0.74	0.007
103345	spectrin alpha 2	0.74	0.032

Affymetrix No.	Cell adhesion genes	Dfl	p value
99842	procollagen, type XIX, alpha 1	0.55	0.025
94117	putative neuronal cell adhesion molecule	0.68	0.0063
93606	immunoglobulin superfamily, member 4	0.74	0.0399

Affymetrix No.	Apoptosis related/cell cycle related genes	Dfl	p value
93921	HLA-B-associated transcript 3	0.53	0.0151
93922	HLA-B-associated transcript 3	0.55	0.0151
95040	programmed cell death 6 interacting protein	0.61	0.0398
103553	homologous to minichromosome maintenance deficient 10	0.7	0.0082
103711	caspase 9	0.71	0.0048
95093	cell division cycle and apoptosis regulator 1	0.71	0.0276
104275	p53	0.74	0.0274
100005	Tnf receptor associated factor 4/ CART1	0.74	0.032

Affymetrix No.	Enzymes	Dfl	p value
103842	DEAD (aspartate-glutamate-alanine-aspartate) box polypeptide	0.42	0.0153
95680	protein phosphatase 1, regulatory (inhibitor) subunit 2	0.47	0.048
95681	protein phosphatase 1, regulatory (inhibitor) subunit 2	0.53	0.0442
97096	protein kinase, cAMP dependent regulatory, type II alpha	0.53	0.0403
103083	Mus musculus hormone-sensitive lipase (Lipe) gene	0.55	0.0452
102639	carbohydrate sulfotransferase 2	0.56	0.0415
104005	beta 1,4- galactosyltransferase, polypeptide 2	0.6	0.0151
104025	thimet oligopeptidase 1	0.61	0.0304
94501	sphingosine-1-phosphate phosphatase 1	0.61	0.0403
100996	Mus musculus putative lysophosphatidic acid acyltransferase	0.62	0.049
94483	casein kinase II, alpha 2, polypeptide	0.63	0.0477
160640	polypeptide N-acetylgalactosaminyltransferase 11	0.66	0.0175
100544	bromodomain containing 7-EST	0.67	0.0336
95054	threonyl-tRNA synthetase	0.68	0.0154
94361	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21 (RNA helicase II/Gu)	0.68	0.0191
99038	adenylosuccinate synthetase 2, non muscle	0.68	0.0442
101964	transketolase	0.69	0.0314
93751	ubiquitin specific protease 19	0.7	0.0397
95161	Nuclear LIM interactor-interacting factor 2	0.71	0.0175
94970	protein phosphatase 4, regulatory subunit 2	0.72	0.0063
100416	maternal embryonic leucine zipper kinase	0.72	0.0326
93752	isoleucine-tRNA synthetase	0.72	0.0472
98079	carbonic anhydrase 14	0.73	0.048
101867	glycerol-3-phosphate acyltransferase gene	0.74	0.0265
93493	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 5	0.75	0.0425

Affymetrix No.	Miscellaneous genes	Df1	p value
96606	calcium binding protein P22	0.51	0.0452
97111	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3	0.54	0.0274
104702	BING4 protein	0.58	0.0065
92782	thymopoietin	0.58	0.0456
160352	poly(rC) binding protein 4	0.6	0.0323
160954	synapsin II	0.6	0.0359
93520	serine/arginine repetitive matrix 1	0.62	0.0279
101894	fused toes	0.63	0.0066
102845	PDZ and LIM domain 7	0.644	0.0478
161884	fragile X mental retardation gene 1, autosomal homolog	0.66	0.0483
98572	DnaJ (Hsp40) homolog, subfamily B, member 11	0.67	0.0403
104106	sno, strawberry notch homolog 1 (Drosophila)	0.68	0.0313
103211	WD repeat and FYVE domain containing 2	0.69	0.025
103096	zinc finger protein 259	0.69	0.0275
103397	HIV-1 Rev binding protein	0.701	0.0472
100974	single stranded DNA binding protein 4	0.71	0.0313
102062	Smarcc1	0.71	0.0423
95232	heterogeneous nuclear ribonucleoprotein L	0.72	0.0439
103953	SEC22 vesicle trafficking protein-like 1 (S. cerevisiae)	0.73	0.0125
97556	acidic (leucine-rich) nuclear phosphoprotein 32 family, member E	0.73	0.032
103789	bromodomain containing 4	0.73	0.0321
93020	reduced expression 3	0.73	0.0477
97254	RNA binding motif protein	0.74	0.0189
95287	LUC7-like 2 (S. cerevisiae)	0.74	0.0326
97251	mitochondrial ribosomal protein S10	0.74	0.0472
160423	mitochondrial ribosomal protein S2	0.75	0.0174

Affymetrix No.	Genes with unknown function	Df1	p value
92664	RIKEN cDNA 2500002G23 gene	0.43	0.0274
96532	ESTs	0.51	0.0425
95151	RIKEN cDNA 2810052M02 gene	0.55	0.0382
160997	cDNA	0.58	0.0352
160804	RIKEN cDNA E130303B06 gene	0.58	0.0442
97666	EST	0.61	0.0279
102135	EST	0.61	0.0477
161756	cDNA	0.61	0.0483
160191	cDNA	0.633	0.0477
94957	cDNA	0.65	0.0066
101732	EST	0.65	0.0379
93301	RIKEN cDNA 1300007B12 gene	0.69	0.0248
95458	RIKEN cDNA 1110001C20 gene	0.69	0.0457
98075	cDNA	0.7	0.0059
97371	RIKEN cDNA 2310022K01 gene	0.71	0.0326
99823	DNA segment, Chr 18, ERATO Doi 232, expressed	0.71	0.0441
160672	expressed sequence AI462438	0.72	0.0279
95028	cDNA	0.72	0.0334
101404	RIKEN cDNA 2310061I09 gene	0.73	0.0274
94549	RIKEN cDNA 1200003O06 gene	0.73	0.0304
96598	RIKEN cDNA D430039C20 gene	0.74	0.0162
95075	RIKEN cDNA 2310006I24 gene	0.74	0.0321
102686	EST	0.74	0.0439
102011	RIKEN cDNA 2610507L03 gene	0.75	0.032
100828	cDNA	0.75	0.0418

4.3.5.1.2 Filtering the upregulated genes: set 2

Figure 4.3.5.2 also illustrates the filtering applied to arrive at the upregulated gene list set 2. An expression cut-off of 1.25 was used to mirror the genes changed in the downregulated set. 870 genes passed this filter which is comparable to the 882 passing the 0.75 filter. For this dataset the non parametric Wilcoxon-Mann-Whitney Test was applied with MTC using BH FDR $p<0.05$. 255 genes passed this test and a further 144 survived with a detection p value <0.1. Table 4.3.5.1.2 lists the genes included in the upregulated gene set 2.

Affymetrix No.	Transcription factors	Dfl	p value
102986	MyoD1	1.53	0.04997
103801	regulatory factor X-associated ankyrin-containing protein	1.46	0.00008
92914	homeo box 2.3; putative; Mouse homeo box (Hox2.3) region.	1.43	0.00016
92564	leucine rich repeat (in FLII) interacting protein 1	1.32	0.00008
99024	Max dimerization protein 4	1.3	0.04997
92195	CCAAT/enhancer binding protein (C/EBP), gamma	1.29	0.04997
97969	nuclear receptor subfamily 1, group H, member 4	1.25	0.04997

Affymetrix No.	Growth factors	Dfl	p value
94222	insulin-like growth factor binding protein 4	1.55	0.00016
95546	insulin-like growth factor 1	1.31	0.00016

Affymetrix No.	Cell cycle regulators	Dfl	p value
162172	neural precursor cell expressed, developmentally down-regulted gene 4	1.67	0.00008
93534	decorin	1.37	0.04997
96603	quiescin Q6	1.33	0.04997
97821	p21 (CDKN1A)-activated kinase 2	1.3	0.00016
104757	ZW10 homolog (Drosophila), centromere/kinetochore protein	1.29	0.04997
100470	mitogen activated protein kinase 10	1.27	0.04997
94971	cyclin-dependent kinase inhibitor 3	1.25	0.04997

Affymetrix No.	Protein signalling pathway genes	Dfl	p value
93672	protein kinase, interferon-inducible double stranded RNA dependent	1.68	0.04997
103394	FXYD domain-containing ion transport regulator 5	1.62	0.04997
101548	synaptojanin 2 binding protein	1.6	0.00029
92805	ARF-like 4 protein; Mus musculus Arl4 gene.	1.54	0.04997
103934	gamma-aminobutyric acid (GABA-A) transporter 3	1.37	0.00029
161609	regulator of G-protein signaling 16	1.32	0.04997
160830	cornichon-like (Drosophila)	1.25	0.04997

Affymetrix No.	Acute phase response genes	Dfl	p value
101575	serine protease inhibitor 1-2	1.63	0.00016
101565	serine protease inhibitor 1-3	1.57	0.00029
101576	serine protease inhibitor 1-2	1.54	0.00016
101572	serine protease inhibitor 1-1	1.52	0.04997
101574	serine protease inhibitor 1-5	1.52	0.04997
93109	serine protease inhibitor 1-4	1.46	0.04997

Affymetrix No.	Proteins increased with stress	Dfl	p value
93573	Mouse gene for Metallothionein-I (three exons).	2	0.04997
101561	metallothionien II; Mouse metallothionein II (MT-II) gene.	1.68	0.04997
162420	selenoprotein K	1.44	0.04997
160280	caveolin, caveolae protein	1.38	0.04997
97821	p21 (CDKN1A)-activated kinase 2	1.3	0.00016

Affymetrix No.	Retinoic acid associated genes	Dfl	p value
95350	transthyretin	1.61	0.00016
102258	stimulated by retinoic acid gene 6	1.41	0.00008

Table 4.3.5.1.2 Upregulated genes. Set 2.

Genes upregulated in the Dfl compared to the wild type cDNA samples and passing the filtering and statistical analysis described in the methods as gene set2. ESTs and genes expressed at 1.25-1.45 wild type levels are not shown in the table to conserve space. **Affy no**: affymetrix number **MA**: relative expression in the Dfl:wild type microarray chips averaged across the 7 arrays.

Affymetrix No.	Haematological proteins	Dfl	p value
103340	Rhesus blood group CE and D	1.9	0.00016
101869	Mouse beta-globin major gene.	1.89	0.04997
162457	hemoglobin alpha, adult chain 1	1.47	0.04997
101703	hemoglobin alpha, pseudogene 3	1.44	0.00016
103535	hemoglobin Y, beta-like embryonic chain	1.33	0.00008
103997	Mouse gene for erythropoietin receptor.	1.32	0.04997
100440	alternatively spliced form of ankyrin; major form of ankyrin found in red blood cells	1.28	0.04997
162479	ferritin light chain 1	1.28	0.04997
101062	heme oxygenase (decycling) 2	1.26	0.04997

Affymetrix No.	Extracellular Matrix Proteins	Dfl	p value
93534	decorin	1.37	0.04997
103356	dedicator of cytokinesis 7	1.32	7.94E-05
161611	Emu1 gene	1.3	0.04997
100569	annexin A2	1.27	0.04997

Affymetrix No.	Cytoskeletal Proteins	Dfl	p value
101578	actin, beta, cytoplasmic	2.55	0.00008
100446	small proline-rich protein 1B	1.55	0.00016
102770	glycophorin A	1.5	0.04997
161612	Tubulin beta-3	1.34	0.04997
104654	actin-like 6	1.26	0.00029

Affymetrix No.	Miscellaneous genes	Dfl	p value
99141	GM2 ganglioside activator protein	2.24	0.00008
101501	imprinted and ancient	1.97	0.00016
160563	elongation factor RNA polymerase II-like 3	1.64	0.04997
102557	Mus musculus histone H2a(A)-613	1.51	0.04997
162201	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2	1.51	0.04997
92480	Zinc finger protein 118	1.51	0.04997
96204	SH3-binding domain glutamic acid-rich protein	1.48	0.00008
94642	guanosine diphosphate (GDP) dissociation inhibitor 2	1.44	0.04997
100702	sex hormone binding globulin	1.42	0.04997
92681	melanoma antigen, family L, 2	1.4	0.00038
93593	epithelial membrane protein 3	1.4	0.04997
162435	fasciculation and elongation protein zeta 1 (zygin I)	1.39	0.04997
92989	Ca<2+>dependent activator protein for secretion	1.39	0.04997
160612	ATP-binding cassette, sub-family G (WHITE), member 1	1.38	0.00016
103955	crystallin, lamda 1	1.36	0.00016
103206	component of oligomeric golgi complex 5	1.35	0.04997
160373	serum deprivation response	1.35	0.04997
161723	WD repeat and FYVE domain containing 2	1.34	0.04997
162275	UPF3 regulator of nonsense transcripts homolog B (yeast)	1.34	0.04997
96935	membrane-associated protein 17	1.33	0.00029
93305	vesicle-associated membrane protein 8	1.32	0.00008
104345	exocyst component protein homolog (S. cerevisiae)	1.32	0.04997
161392	leucine rich protein, B7 gene	1.31	0.04997
102943	autophagy 7-like (S. cerevisiae)	1.3	0.04997
92919	5-hydroxytryptamine (serotonin) receptor 3A	1.3	0.04997
92439	otoconial matrix protein; Mus musculus otoconin-90 (Oc90) gene, partial cds.	1.29	0.00029
96071	ribosomal protein S25	1.29	0.04997
162426	secretory carrier membrane protein 1	1.27	0.04997
92725	growth arrest specific 8	1.27	0.04997
93018	interferon induced transmembrane protein 3-like	1.25	0.04997

Affymetrix No.	Immune related genes	Dfl	p value
98543	cathepsin S	1.74	0.00016
102250	T cell cytokine receptor	1.63	0.00008
101728	complement component 5, receptor 1	1.44	0.00038
104239	pan hematopoietic expression	1.4	0.04997
103570	C1q and tumor necrosis factor related protein 3	1.36	0.04997
101516	CD59a antigen	1.3	0.04997
100397	TYRO protein tyrosine kinase binding protein	1.29	0.04997

Table 4.3.5.1.2 Upregulated genes. Set 2.

Affymetrix No.	Enzymes	Dfl	p value
161763	phosphatidylinositol-4-phosphate 5-kinase, type II, gamma	1.84	0.00038
96038	ribonuclease, RNase A family 4	1.66	0.04997
162129	protein tyrosine phosphatase, receptor-type, F interacting protein, binding protein 2	1.48	0.04997
102797	retinal short-chain dehydrogenase/reductase 1	1.45	0.04997
160122	asparaginase like 1	1.45	0.04997
96258	microsomal glutathione S-transferase 3	1.43	0.04997
160577	ATPase, class II, type 9A	1.41	0.04997
161997	aldehyde dehydrogenase 2, mitochondrial	1.41	0.04997
103038	guanylate cyclase activator 1a (retina)	1.4	0.00016
100425	spleen tyrosine kinase	1.4	0.00029
103573	phosphatidylinositol-4-phosphate 5-kinase, type 1 alpha	1.4	0.04997
96287	deoxyuridine triphosphatase	1.4	0.04997
97496	protein kinase C, delta binding protein	1.34	0.04997
160746	Rho-associated coiled-coil forming kinase 1	1.34	0.04997
93009	glutathione S-transferase, mu 2	1.33	0.04997
102035	thiopurine methyltransferase	1.31	0.00016
160194	glutaryl-Coenzyme A dehydrogenase	1.31	0.00016
104509	cholesterol 25-hydroxylase	1.31	0.04997
162343	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit	1.31	0.04997
98859	Mus musculus (clone lambda-MG5.3) acid phosphatase type 5 gene, complete cds.	1.31	0.04997
97821	p21 (CDKN1A)-activated kinase 2	1.3	0.00016
161337	tryptophanyl-tRNA synthetase	1.3	0.00038
160085	thiosulfate sulfurtransferase, mitochondrial	1.3	0.04997
101991	flavin containing monooxygenase 1	1.29	0.04997
92441	fibroblast activation protein	1.29	0.04997
160853	phosphatidylserine synthase 2	1.28	0.00008
103021	Mitogen activated protein kinase kinase kinase 1	1.26	0.04997
95590	asparagine-linked glycosylation 5 homolog	1.25	0.04997

Affymetrix No.	Genes with unknown function	Dfl	p value
102233	RIKEN cDNA 4833442J19 gene	1.71	0.00016
93155	RIKEN cDNA 1500032A09 gene	1.68	0.04997
160250	RIKEN cDNA 1810029G24 gene	1.63	0.04997
161464	RIKEN cDNA 1810013H02 gene	1.57	0.00008
104694	similar to zinc finger protein 40	1.56	0.00008
95989	ESTs	1.49	0.04997
98458	RIKEN cDNA 1190007F08 gene	1.48	0.04997
100547	RIKEN cDNA 4921506I22 gene	1.43	0.04997
96122	RIKEN cDNA 2310016A09 gene	1.43	0.04997
161134	RIKEN cDNA 2610201A13 gene	1.4	0.00016
162258	RIKEN cDNA 2410012A13 gene	1.4	0.00029
101595	ESTs	1.34	0.04997
97874	RIKEN cDNA 1500032D16 gene	1.32	0.00038
95442	hypothetical protein MGC18837	1.31	0.00016
160486	RIKEN cDNA 1810010N17 gene	1.3	0.00008
101697	cDNA	1.3	0.00016
104726	RIKEN cDNA 1700013L23 gene	1.3	0.04997
97102	Transcribed sequences	1.3	0.04997
160212	RIKEN cDNA 2810002P21 gene	1.29	0.04997
98051	expressed sequence AI114950	1.29	0.04997
160831	expressed sequence AI838661	1.28	0.00038
161627	cDNA	1.28	0.04997
160465	RIKEN cDNA 3300001H21 gene	1.27	0.00029
161424	RIKEN cDNA 1110049F12 gene	1.27	0.04997
160387	RIKEN cDNA 1110055L24 gene	1.26	0.04997
94389	RIKEN cDNA 2310034K10 gene	1.26	0.04997
160600	RIKEN cDNA 1700009P03 gene	1.25	0.04997
160717	RIKEN cDNA 1110059E24 gene	1.25	0.04997

Table 4.3.5.1.2 Upregulated genes. Set 2.

4.3.5.2 Microarray results of the genes within the Df1 deleted region

Table 4.3.5.2a and b illustrates the microarray results for the genes within the *Df1* deleted region. 13 of 22 genes mapping within the *Df1* deletion or at its endpoints are present on the MG-U74a array. *Dgcr6*, *Scl25a1* and *Hft9c* are each represented by two oligonucleotide probe sets. All but two of the genes (*Gp1bβ* and the 2nd probe for *Hft9c*) were detected at above the threshold control strength used in the data filtering process described in the materials and methods. *Gp1bβ* is expressed in platelets and would not be expected to give an above detection threshold signal in E10.5 embryos. Nine *Df1* genes were identified as significantly downregulated and were included in set 1a, with *T10* and *Ufd1l* also reduced in expression (relative expression of 0.79 and 0.63 respectively) but failed the statistical filtering (p values of 0.178 and 0.202 respectively). However the relative expression of *Es2* to wild type levels was 1.20.

Probe sets for *Df1* genes not detected as downregulated in the filtering described above were examined more closely for reasons for lack of consistency. This analysis was performed by inspecting the probe sequence for the genes available on the Affymetrix website (https://www.affymetrix.com). Details of this analysis are shown in table 4.3.5.2c. The NCBI Blast facility (http://www.ncbi.nlm.nih.gov/BLAST/) was employed to check the sensitivity of the 16 oligonucleotides used for probes for each gene. Sequence inspection revealed that *Htf9c* (second probe) and *Ufd1l* had poor resemblance to the published gene, whereas the match for *Es2* and *T10* probes was 100%. *Es2* and *Ufd1L* are located at the 5' and 3' boundaries of the *Df1* targeted region respectively. As the promoter region of both of these genes is deleted in the *Df1* mouse, expression of these genes is not expected. Both of these genes were followed up by RTQPCR (see Chapter 5).

Table4.3.5.2a

Affy No.	Df1 deleted genes-statistically significantly downregulated	MA	p value	SD	Signal
102403	cell division cycle 45 homolog (S. cerevisiae)-like (CDC45)	0.48	0.0053	0.15	V.High
104516	claudin 5	0.51	0.0097	0.18	High
98573	RAN binding protein 1 (ranbp1)	0.54	0.0058	0.14	V.V.High
96337	septin 5	0.54	0.0139	0.2	Medium
94956	DiGeorge syndrome critical region gene 6 (DGCR6)	0.54	0.0336	0.29	High
160437	thioredoxin reductase 2 (TXNRD2)	0.58	0.0275	0.24	Medium
100525	HpaIIItiny fragments (Htf9c)	0.59	0.0048	0.11	V.High
94807	Scl25a1	0.59	0.0048	0.11	V.High
98535	catechol-O-methyltransferase (COMT)	0.59	0.0048	0.11	High
162358	Scl25a1	0.64	0.0273	0.2	High

Table4.3.5.2b

Affy No.	Df1 deleted genes-not passing statistical testing	MA	p value	SD	Signal
93303	Ubiquitin dependent proteolytic protein (Ufd1l)	0.63	0.202	0.38	High
95025	DNA segment, Chr 16, human D22S680E, expressed (T10)	0.79	0.178	0.25	High
160649	Platelet glycoprotein 1b beta (Gp1bβ)	1.02	0.9564	0.46	Low
162149	HpaIIItiny fragments (Htf9c)	1.02	0.9436	0.62	Low
160574	Expressed sequence 2 embryonic lethal (Es2)	1.2	0.2928	0.61	Medium
161512	DiGeorge syndrome critical region gene 6 (DGCR6)	1.3	0.2486	0.74	High

Table 4.3.5.2c

Df1 Gene	Htf9c	DGCR6	UFD1L	ES2	Gb1bβ	T10
Probe Inspection	Short sequence	Short sequence	Short sequence	Good match	Good match	Good match
Two way BLAST	No alignment	No alignment	No alignment	Good alignment	Good alignment	Good alignment
General BLAST	No match	No match	No match	Good match	Good match	Good match
Probe sequence derived from 3' UTR?	N/A	Yes	Yes	No	No	Yes
Expressed at E10.5?	No information	Yes	Yes	Yes	No	No information
Other				At centromeric Df1 deletion boundary		

Table 4.3.5.2 Genes within the Df1 deleted region

a) Details of the genes within the Df1 deleted region on mouse chromosome 16 that passed the filtering and statistical analysis described in the methods for gene set1a.
b) Genes within the Df1 deletion that did not pass the filtering and statistical testing described in the methods for gene set1a.
c) Details of the Affymetrix probe analysis for the Df1 genes that didn't pass the statistical analysis. The probe sequences were assessed by inspection, 2 way BLAST alignment and a general BLAST analysis to assess their homology with the gene they are representing.
Abbreviations:
Affy no: affymetrix number given to each gene (full information available at www.affymetrix.com).
MA: relative expression in the Df1: wild type microarray chips averaged across the 7 arrays. p values were arrived at after statistical analysis described in the methods for gene set1a. SD: standard deviation of the relative changes. Signal: raw signal intensity values in arbitary units (low:<10, Medium:10-50, High: 50-100, V.High:100-200, V.V high: >200).

4.3.5.3 Microarray results of the genes near to the Df1 deletion endpoints

Table 4.3.5.3 gives details of the genes outside but near the *Df1* deleted region. Microarray analysis of *Hira* and *Nlvcf (Mrp140)*, genes just distal to the *Df1* deletion, revealed expression levels of 1.02 and 0.95 respectively in *Df1* versus wild type embryos. Expression levels of *Stk22a* and *Stk22b*, genes just proximal to the deletion, were too low to be considered reliable, as expected for testes-specific kinases. There were two probe sets for *Dgcr2*, the next closest gene proximal to the deletion, and both were expressed at relatively high levels and neither showing alteration of expression level. Database interrogation revealed that the probes for these genes were 100% matches with database entries. Therefore, there was no evidence that the *Df1* deletion had affected neighbouring genes.

Crkl is located approximately 400kb proximal to *Es2* and it has been proposed that deletion of *CRKL* may contribute to the human syndrome (Guris et al. 2001). Interrogation of the microarrays showed a relative expression of 0.77 for *Crkl* in *Df1* versus wild type samples, which failed the selection criteria for significantly changed genes as it is just above the cut-off of <0.75. This gene was followed up by RTQPCR (Chapter 5).

Table 4.3.5.3

Details of the genes outside but near the Df1 deleted region on mouse chromosome 16.

Affy No.	Genes outside but near the Df1 deleted region	MA	p value	SD	Signal
103848	v-crk sarcoma virus CT10 oncogene homolog (avian)-like (CRKL)	0.77	0.0834	0.61	High
103360	serine/threonine kinase 22A (spermiogenesis associated) (Stk22a)	1.73	0.0475	0.64	Low
102846	serine/threonine kinase 22B (spermiogenesis associated) (Stk22b)	1.79	0.1429	1.48	Low
100142	DiGeorge syndrome critical region gene 2 (DGCR2)	0.94	0.763	0.42	High
161247	DiGeorge syndrome critical region gene 2 (DGCR2)	1.02	0.7297	0.15	High
103009	histone cell cycle regulation defective homolog A (S. cerevisiae; HIRA)	1.02	0.2422	0.74	High
93603	mitochondrial ribosomal protein L40 (Mrp140)/ NLVCF	0.95	0.7816	0.46	V.High

See Table 4.3.5.2 (previous page) for a description of the abbreviations.

4.4 Discussion

4.4.1 What is meant by a gene "target"?

There are several reasons why a gene might be highlighted as significantly changed in a microarray experiment. Firstly the change might be an artefact. For instance the probes used on the chip might be non-specific for the target, or the hybridisation might be of low intensity and give a low signal. It is anticipated that the majority of these artefactual changes would be eliminated by the filtering applied to the raw dataset of differentially regulated genes. Genes with low signal intensity values were excluded, and only genes that showed consistent changes that passed stringent statistical tests were considered further as possible direct or indirect targets of the *Df1* deletion. Also the gene could be significantly changed due to a tissue loss artefact. This is more commonly seen in microarray analysis using homozygous systems.

Secondly the gene could be an indirect target of a gene within the *Df1* deletion. It is likely that the majority of genes that are detected as changed in a microarray experiment are indirect targets. Indirect targets are easy to speculate on but difficult to prove. Obviously the network of genetic changes are identified at a snapshot in time whereas most of the genes will be in a state of flux. Thousands of genes will be influencing each other to try and maintain survival of the embryo and homeostasis.

Thirdly the gene could be a direct target of the deletion. Such targets are more useful in delineating genetic pathways and the aetiology of disease, but are also difficult to prove without sophisticated molecular techniques.

4.4.2 Sensitivity of Affymetrix microarray technology to detect heterozygous expression changes.

As the *Df1* mouse is haploinsufficient for 22 genes, the expression of these genes is expected to be approximately 50% of their values in the wild type embryos. This has led to the use of these genes as controls in interpreting the validity of the microarrays. Nine out of 13 (69%) of the genes deleted within the *Df1* deletion are reduced in expression level on the *Df1* microarray chips compared to the wild type chips and passed the stringent statistical testing applied to analyse the data. A further two genes (*T10* and *UFD1L*) were downregulated but not to the extent that they passed statistical testing (4.3.5.2b) and most of the *Df1* genes that didn't appear downregulated had a valid

reason (table 4.3.5.2c). A recent paper has highlighted the inaccuracies of Affymetrix probe design for some of the older Genechips such as MG-U73Av (Dai et al. 2005). The paper points out that the earlier Genechips probe selection relied on earlier genome and transcriptome annotation which is different from current knowledge. This clearly has implication for all genes followed up as potential "targets" and demonstrates the importance of verification methods.

The results for the *Df1* genes signifies that the methodology and filtering employed in producing the downregulated dataset is sensitive enough to detect approximately two-fold changes. The results add confidence to the assessment of other genes passing the statistical analysis as also being differentially expressed.

4.4.2.1 Experimental limitations

The range of relative expression of *Df1* deleted genes was between 0.48 and 0.64 in *Df1* arrays. This is not surprising given these genes are haploinsufficient in every tissue type in the branchial arch dissection. However it is likely that genes that are regulated by the *Df1* genes may not show uniform changes in all tissues, and therefore any potential "targets" genes are probably diluted by including all of the different tissue types within the branchial arch region. It is therefore possible that heterozygous changes are at the limit of detection by microarray, making validation experiments crucial. This concept is further discussed in Chapter 5.

Inevitably in this type of experiment there will be false positive and false negative differentially expressed genes passing the statistical filtering. For this reason the genes that are candidates for true differential expression have been followed up by RTQPCR and *in situ* hybridisation (Chapter 5). One example of a false negative gene is *Fgf8*. There is significant experimental evidence in the literature that *Fgf8* is a downstream target of *Tbx1* (Abu-Issa et al. 2002; Vitelli et al. 2002b; Brown et al. 2004). However despite the gene being present on the MG-U74Av chip this gene didn't appear downregulated. There are several possible reasons for this. One is poor probe design or probe cross hybridisation. There are two probe sets for *Fgf8* (Affy no 97742 and no 170994). The probe for Affymetrix no 97742 has 100% identity with *Fgf8* on BLAST analysis. The second probe for Affymetrix no 170994 does not show any identity to *Fgf8* on BLAST analysis. A second possible reason is that although *Fgf8* is a recognised

Tbx1 target, it may not be a downstream target in all the tissues used in the branchial arch dissection and also other gene within the *Df1* deletion could alter its expression level. Thirdly the evidence for *Fgf8* as a *Tbx1* target is primarily in *Tbx1* homozygous mouse embryos (Vitelli et al. 2002).

4.4.2.2 Identification of genes represented by more than one probe in gene set 1a

It is noteworthy that some genes are represented by more than one probe set. Where more than one probe set has passed the statistical filtering adds confidence that this gene is "true positive". There are four examples of such genes all of which passed filtering and statistics to be in set 1a. These genes are detailed in table 4.4.2.2 below:

Gene	Affymetrix probe no. 1	Affymetrix probe no.2
SRY-box containing gene 3	93921	93922
HLA-B-associated transcript 3	93921	93922
Protein phosphatase 1, regulatory (inhibitory) subunit 2	95680	95681
Guanine nucleotide binding protein, alpha 12	97227	97226

Table 4.4.2.2 Genes represented by two Affymetrix probe sets

4.4.2.3 Detection of two genes that are now known Tbx1 targets.

In addition to the detection of 69% of the *Df1* deleted genes by this methodology, two further genes have recently been described as *Tbx1* targets. *Foxa2* is a gene that is present in dataset 1b (its detection p value was too high to warrant its inclusion in set 1a). This is a gene that is known to be upstream of *Tbx1* and more recently has been postulated to exist in a feedback loop where it is also downstream of *Tbx1* (Hu et al. 2004).

Crkl is a gene that was only downregulated to 0.77 on the microarray and therefore was not included in dataset 1a (where a cut off of <0.75 was used). Any selected array data are finite sets by definition. However Genespring makes it possible to interrogate the dataset to look for levels of genes of interest. *Crkl* is one such gene as the homozygous mouse has a phenotype similar to 22q11DS (Guris et al. 2001) and some 22q11DS patients exist with deletions encompassing *CRKL* and not *TBX1* (figure 1.4.2.2) (Kurahashi et al. 1996; O'Donnell et al. 1997; McQuade et al. 1999; Garcia-Minaur et

al. 2002). Recent work has described *Tbx1/Crkl* mouse transheterozygotes as being more severely affected than the heterozygous mice alone, signifying an epistatic interaction between the two genes. Further work by this group has shown that *Crkl* is downstream of *Tbx1* (Guris et al. 2006).This gene is followed up and discussed in Chapter 5.

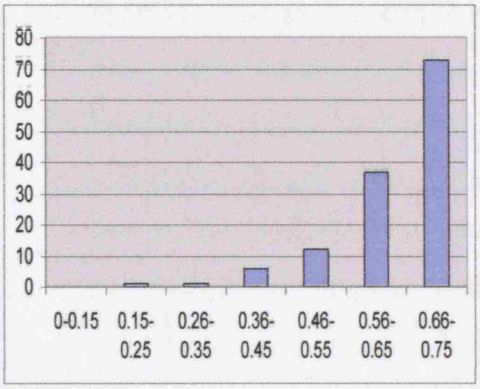
4.4.2.4 Detection of genes that regulate each other

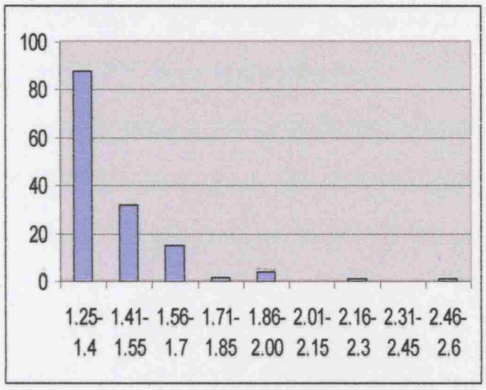
Of note is the identification of a gene pair, one of which regulates the other. The gene *protein-kinase, interferon-inducible double-stranded RNA dependent inhibitor* (*PRKRIR*: Affymetrix number 99975) from set 1a (table 4.3.5.1.1) that was downregulated to 0.63 of wild type levels in *Df1* embryos: has been shown to inhibit the gene *protein kinase, interferon-inducible double stranded RNA dependent* (*PRKR*: Affymetrix number 93672), which is upregulated 1.7 fold and is in gene set 2 (table 4.3.5.1.2) (Gale, Jr. et al. 1998). This adds weight to the value of these genes as possible true positives and enhances the evidence that the statistics and filtering applied is valid.

4.4.3 Patterns of gene changes

From figure 4.3.5.2 it can be seen that generally the numbers of genes in the up and down regulated lists that pass through each stage of filtering are similar. It is interesting to note that the statistical testing used for these two datasets was slightly different. Set 1a was arrived at using a parametric t-test, whereas a non parametric test was used for the upregulated gene set 2 (zero genes passed the testing for set 2 if the parametric test was used). To analyse the reasons for this in more detail the up and down regulated datasets are represented in a bar chart format to inspect the distribution of changed genes (figure 4.4.3).

Figure 4.4.3 The distribution of gene changes in the gene set 1a and set 2. X axis: number of genes, Y axis: relative categories for downregulation or upregulation.





Set 1a Downregulated gene distribution Set 2 Upregulated gene distribution

From the graphs above it can be seen that set 1a follows a more even or normal distribution, whereas set 2 is more unevenly distributed. It is possible this is why the parametric testing could not be applied to set 2: because the gene changes do not follow a "normal distribution".

4.4.4 Follow-up work

Lists of genes from microarray experiments have limited meaning until further work is done to verify the changes seen. Generally there several avenues of verification:

1) Bioinformatic techniques can initially be employed to examine data lists and find genes that are biologically relevant because they are known to act on certain pathways or have certain expression patterns. Fortunately the Affymetrix website (www.affymetrix.com), allows quick and easy gene analysis by providing links to NCBI (http://www.ncbi.nlm.nih.gov/) and other bioinformatic websites. Candidates for further verification were initially examined in this way.
2) Alternative RNA techniques such as RTQPCR can be used to examine the degree and nature of differential gene expression. This is important to provide some quantification of changes, whereas *in situ* hybridisation techniques can be employed to examine qualitative information on anatomical regions of possible change. The latter technique is more sensitive for large gene changes.
3) Identification of potential direct transcriptional targets using bioinformatics techniques can provide some information on promoter regions of genes. *In vitro* promoter assays may support these findings.

4) Expression changes do not always translate into changes in protein levels. If there is an abundance of tissue then protein investigations such as Western hybridisations are a useful avenue of verification. In the case of this study there was insufficient protein in the mouse embryo for Western analysis.
5) Ultimately *in vivo* models can be employed to assess the functional relevance of a potential "target" gene. By crossing mice heterozygous for the two genes of interest it is possible to analyse the affect on the phenotype to see if the genetic pathways of the two genes overlap.

Chapter 5 Verification of *Df1* microarray results

5.1 Introduction

In paragraph 4.4.4 some of the methods of microarray target verification were briefly discussed from further RNA work to functional studies. Protein analysis would have been difficult in this project due to constraints on tissue abundance. Functional analysis would be the gold standard method of follow-up, but before this is possible further RNA work is the simplest and most efficient starting point. This chapter describes the results of the two main methods of follow-up employed in this project: real time quantitative PCR (RTQPCR) and whole mount *in situ* hybridisation.

5.1.1 Real Time Quantitative PCR

RTQPCR is an increasingly popular method used to assess mRNA quantification, allowing relative and absolute analysis of gene expression from low concentrations of starting template. It is widely used for validation of expression data produced from microarray experiments (Nadon & Shoemaker 2002) as it sensitive, allows high sample throughput and has a broad range of starting concentrations. As RNA cannot serve as a template for PCR, the first step is the reverse transcription of the RNA template into cDNA. This is then followed by the RTQPCR reaction itself which is an exponential amplification of the cDNA in a PCR reaction.

5.1.1.1 Chemistry

RTQPCR using the intercalating dye SYBR Green utilises the incorporation of the fluorescent dye in a PCR reaction mixture (figure 5.1.1.1.1). SYBR green binds to double-stranded DNA and gives off a fluorescent signal proportional to the DNA concentration during the exponential phase of the PCR reaction (figure 5.1.1.1.2). Reactions may be analysed and compared whilst still in the linear phase of exponential amplification before any plateau phase occurs. Templates with a high initial starting concentration will demonstrate an earlier exponential rise in fluorescence than a more dilute sample. In order to analyse an amplification plot a threshold is arbitrarily assigned to where all the samples are in their linear phase. The threshold cycle (Ct) is the PCR cycle at which each amplification plot reaches this threshold (figure 5.1.1.1.2).

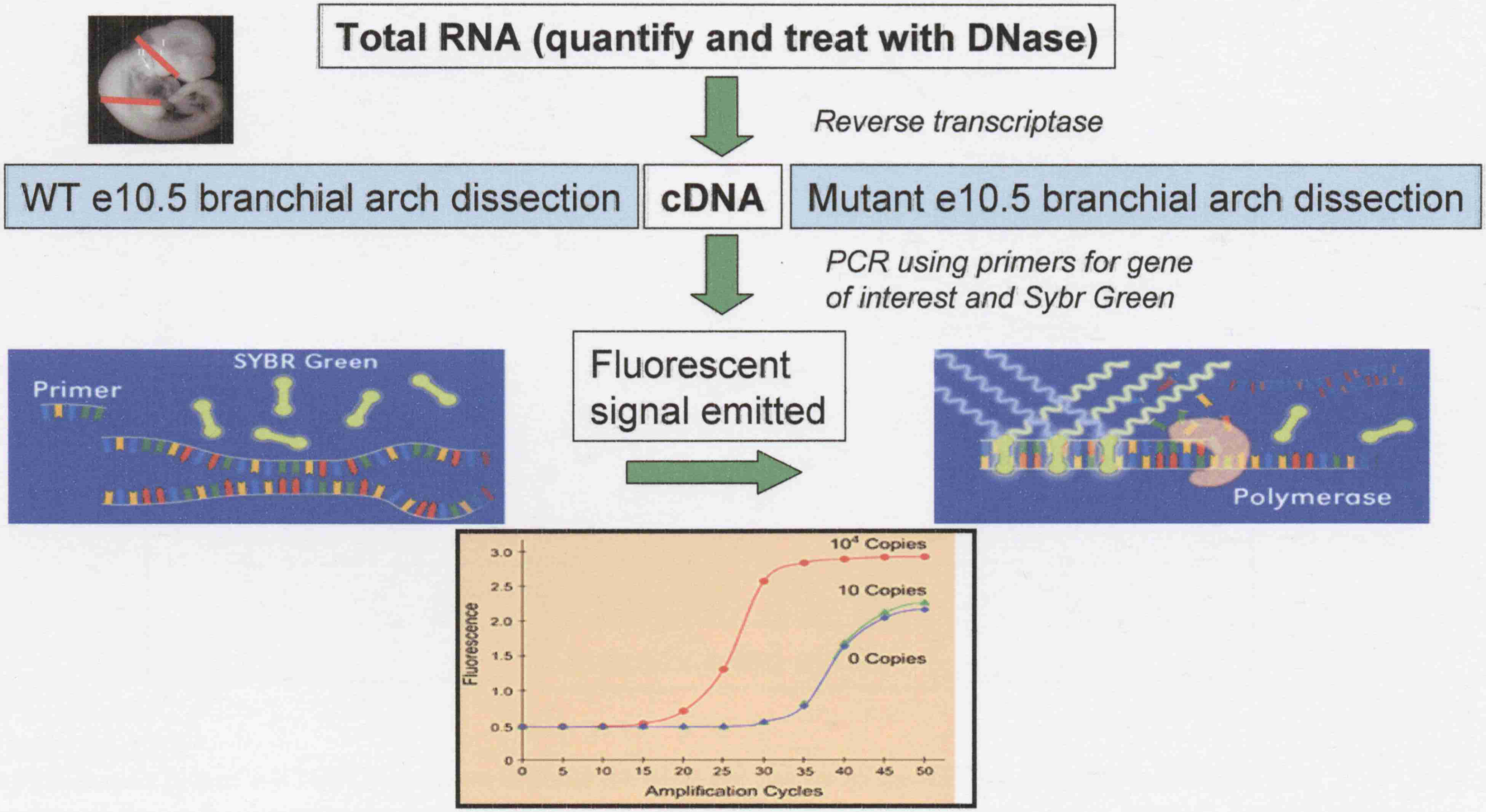


Figure 5.1.1.1.1

Steps involved in RTQPCR analysis to quantify total RNA.

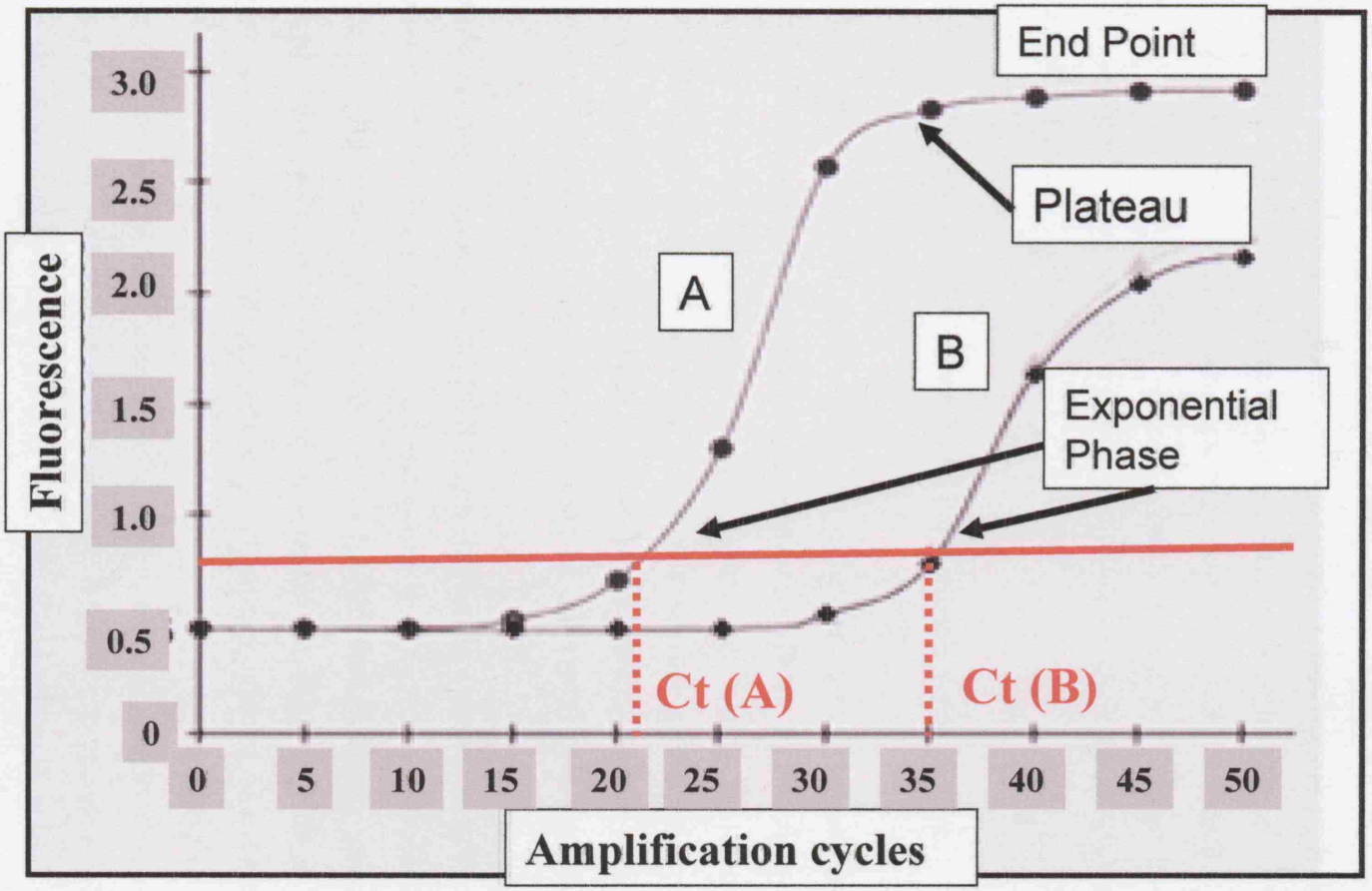


Figure 5.1.1.1.2

Amplification plot of a RTQPCR reaction to show the different phases of the reaction in samples A and B. Red horizontal line is the fluorescence threshold and Ct (A) and Ct (B) are the threshold cycles for samples A and B where the vertical red dotted lines cross the X axis.

5.1.1.2 Technology

All real time PCR technologies aim to minimise variability between samples and monitor fluorescence signals as they are generated, recording a threshold cycle where fluorescence rises above background but before the reaction reaches a plateau. There are various different types of instrumentation available all with different advantages and disadvantages. They all work along a similar principle usually consisting of a thermal cycler with laser, optics and a detector system (typically a charge coupled device camera) enabling single sample imaging on a cycle-by cycle basis. This technology is illustrated in figure 5.1.1.2.

5.1.1.3 Absolute and relative quantification and sample normalisation.

Quantification of mRNA transcription can either be relative or absolute. Absolute quantification determines the input copy number of cDNA transcripts by relating the PCR signal to a standard curve comprising of serial dilutions of known concentrations of cDNA. A standard curve has to be constructed for each primer pair used, to ensure accurate reverse transcription and PCR amplification profiles. Absolute quantification is not always necessary for the purposes of assessing differential gene expression.

Relative quantification relates the PCR signal of a test to a control sample. To allow for differences in initial sample input, it is necessary to correct for differences by using an internal control which is typically an unchanged housekeeping gene. The RNA encoding *Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)* is a ubiquitously expressed, moderately abundant message. It is frequently used as an endogenous control for RTQPCR analysis because in some experimental systems its expression is constant at different times and after experimental manipulation (Edwards & Denhardt 1985; Winer et al. 1999). There are several methods of relative quantification calculation, but one of the more widely used methods is the "$2^{-\Delta\Delta Ct}$" method of data analyses (Livak & Schmittgen 2001). The method uses the assumption that the PCR efficiency for each primer pair is equal to one. This assumption is tested by comparing standard curves of target and control amplicons.

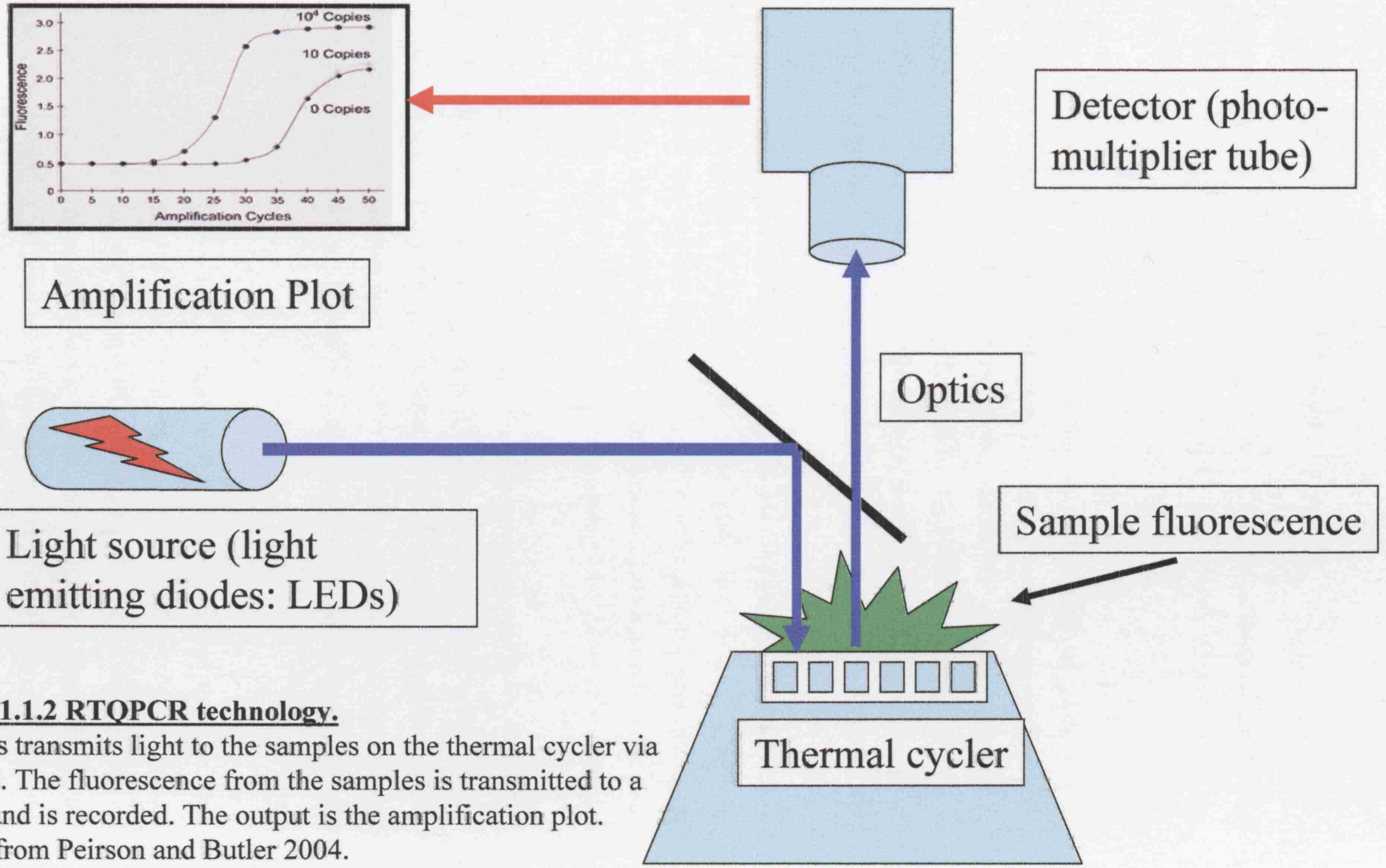


Figure 5.1.1.2 RTQPCR technology.
The LEDs transmits light to the samples on the thermal cycler via the optics. The fluorescence from the samples is transmitted to a detector and is recorded. The output is the amplification plot. Adapted from Peirson and Butler 2004.

5.1.2 Whole Mount In Situ Hybridisation

This technique is used to detect and locate synthesis of specific mRNAs in tissues. It is based on the complementary pairing of labelled RNA probes with RNA *in situ* in an embryo for instance.

Traditionally this technique has been performed on sectioned tissue using probes labelled with radioactive markers. Photographic film emulsion was used for the detection of the signal. Non-radiolabelled probes allow the *in situ* hybridisation procedure to be used on whole embryos (Tautz & Pfeifle 1989). This technique is very sensitive with high resolution and the use of single-stranded complementary RNA probes (riboprobes) has improved the sensitivity even further (Lehmann & Tautz 1994).

In the original method DNA fragments were used that were labelled by random priming with DIG-dUTP. Since RNA-RNA hybrids are more stable than RNA-DNA hybrids, elevated hybridisation temperatures can be employed, which result in higher specificity and less background. The detection of the hybridisation signal is usually done with chromogenic substrates which develop a colour at the place where the probe has bound.

RNA probes are prepared by *in vitro* transcription of linearised plasmid DNA with RNA polymerase. Plasmid vectors containing the polymerase promoters from bacteriophages T3, T7 or SP6 are used. The length of the RNA is determined by the position of the DNA restriction site, downstream from the RNA promoter. In addition to synthesizing an antisense probe it is usual to synthesize a sense probe to use as a control for non-specific binding.

Riboprobes are often labelled with Digoxigenin (Dig) which is linked to the uridine nucleotide which is incorporated into the riboprobe during the *in vitro* transcription reaction. This probe can then be detected with antibodies conjugated to a number of different labels such as alkaline phosphatase, which results in a blue precipitate when the enzyme is incubated in the presence of the substrate NBT/BCIP (Tetrazolium salt/ 5-bromo-4, chloro-3, idolyl-phosphate).

5.1.3 Selection of genes to follow-up

There are limits to the number of genes that are practical to follow-up in a microarray screen. It is anticipated that the use of seven array replicates for wild type and *Df1* RNA has reduced the number of false positive genes to a minimum. Fairly stringent statistical filtering to the gene lists was used to reduce the genes to a manageable number. For each gene it was useful to inspect the raw signal intensity data available from Genespring. Attention was focused on genes that were biologically relevant, but also had a raw data profile that was reliable.

A thorough literature search on all of the genes in the list was the first initial step to identify information on the expression pattern and function of each gene. Sources of information were mainly NCBI (http://www.ncbi.nlm.nih.gov.) and the Affymetrix (https://www.affymetrix.com) website.

5.1.4 Aims of verification

The aims of target verification were initially to answer the questions posed in paragraph 4.1.3.3. In addition I wanted to examine the change in relative gene expression across time so I opted to verify targets by using a "somite series". Embryos were pooled from each somite stage from 16 up to 36 somites which represent e9.0 to e10.75 embryos respectively These stages were selected as they represent the time of formation of the pharyngeal arch arches. The relative expression of *FoxA2/Shh* and *Tbx1* across the somite series was studied as these genes form part of a biological pathway (see discussion).

An additional question to be answered was whether the genes from the *Df1* microarray screen were targets of the *Df1* deletion genes in general or more specifically targets of *Tbx1*. This was investigated by performing RTQPCR using E10.5 *Tbx1* heterozygous embryo cDNA as template.

5.2 Experimental Methodology

For a more detailed description of the methodology refer to chapter 2.

5.2.1 RTQPCR

An overview of the methodology employed for RTQPCR is shown in figure 5.1.1.1.1.

5.2.1.1 Template preparation

Branchial arch dissections were performed on embryos from 18-36 somites and embryos were genotyped using yolk sac DNA. Embryos were separated into *Df1* and wild type and pools of between 3 and 6 embryos were made for 9 different somite stages. RNA was extracted using TRIzol and the embryos were treated with DNase to eliminate genomic double-stranded DNA which would interfere with the RTQPCR reaction. RNA was quantified and reverse transcribed into cDNA.

In addition E10.5 mouse embryos from *Tbx1/+ vs* wild type crosses were dissected and RNA was extracted as described. The embryos were genotyped using yolk sac DNA and pooled into 4 embryos each of *Tbx1* heterozygotes and wild types. The RNA was treated with DNase and reverse transcribed into cDNA.

In both cases random hexamer primers were used in the reverse transcription PCR reaction. The alternative is an oligo(dT) primer which binds to the poly A tail of the RNA and reverse transcribes from this end. The latter method is not recommended for large genes as identification of transcripts at the 5' end of the gene may be unreliable.

5.2.1.2 Primer design

To maximise gene specificity, primers were mainly designed such that they annealed to the 5' end of the last exon and to the 3' UTR (untranslated region). Primer sequences were also analysed by the NCBI BLAST facility to minimise cross-hybridisation to other genes. Particular care was taken, where possible, to design primers that annealed to sequences in exons on both sides of an intron to avoid the amplification of contaminating genomic DNA. Amplicon length was mainly restricted to between 80

and 150 base pairs as recommended in the SYBR green protocol (QIAgen). Primer specificity was maximised by ensuring primers were at least 20 base pairs long and by avoiding a T at the 3' end (primers with a T. nucleotide at the 3' end have a greater tolerance of mismatch).

One of the disadvantages of using SYBR green for RTQPCR is primer-dimer formation. This can be avoided by ensuring the primers are designed such that they have a maximum GC content of 40% for the five last 3' end nucleotides and by avoiding complementary sequences within a primer sequence and between the primer pair (Vandesompele et al. 2002). Use of SYBR green has the advantage that PCR amplicons and/or non-specific amplification products can accurately be distinguished by the generation of DNA melting curves and first derivative melting peaks (Ririe et al. 1997). These measures were used for the primers designed in this project.

5.2.1.3 PCR reaction set-up

cDNA PCR reactions were mainly set up on a 96-well plate. Primers aliquots were subjected to 3000 Joules from the UV Stratalinker and cDNA was always diluted with fresh milliQ water to minimise double-stranded DNA contamination. Diluted cDNA samples were distributed into wells using a multipipette to minimise pipetting errors. Triplicate reactions were always used for each gene to reduce inter-sample variation. *GAPDH* control experiments were always run on the same plate the test genes.

5.2.1.4 Melt curve assessment

Performing a melt curve analysis can aid in product identification and determination of product homogeneity (Ririe et al. 1997). The principle behind melt curve analysis is that as the fluorescence of SYBR green is dependent on annealing, a decrease in fluorescence is obtained as melting progresses. To carry out a melting curve analysis, the temperature is increased very slowly from a low temperature (e.g. 65°C) to a high temperature (e.g. 95°C). At the lower temperature, all PCR products are double stranded, so SYBR Green binds to them and fluorescence is high, whereas at higher temperatures, PCR products are denatured, resulting in rapid decreases in fluorescence. The fluorescence is measured continuously as the temperature is increased and plotted against temperature. A curve is produced because fluorescence decreases slightly

through the lower end of the temperature range, but decreases much more rapidly at higher temperatures as the melting temperatures of non-specific and specific PCR products are reached. The detection systems calculate the first derivatives of the curves (temporal change in fluorescence), resulting in curves with peaks at the respective specific melting temperature (figure 5.2.1.4).

Melt curve analysis is important to look for primer-dimer formation, and was performed for all of the RTQPCR reactions. Curves with peaks at a temperature lower than that of the specific PCR product indicated the formation of primer-dimers (figure 5.2.1.4). If primer-dimer formation occurs then these will contribute to the fluorescent increases recorded for the reaction and give misleading results. All primers showing primer-dimer formation on melt curve analysis were redesigned.

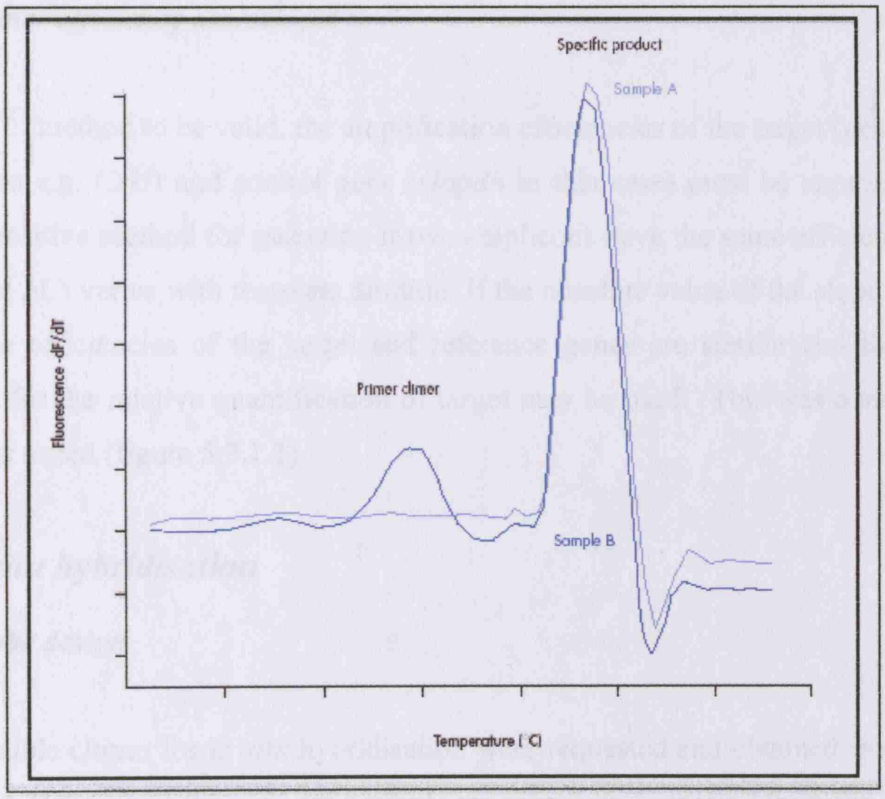


Figure 5.2.1.4

Melting curve analysis of two samples (A and B). Sample A yields only one peak resulting from the sepcific amplificaiton product (primer dimers not co-amplified). Sample B shows a peak from the specific product and a peak at a lower temperature from the amplificaiton of primer-dimers.

5.2.1.5 Normalisation

It is necessary to use a control gene in RTQPCR reactions to enable a correction to be made for the amount of initial cDNA template. The microarray data was inspected for a "housekeeping" gene that was not significantly different between the *Dfl* and the wild type samples. *β-Actin* was excluded as a control as it appeared to be "upregulated" in the *Dfl* sample set compared to wild types (see table 4.3.5.1.2). *Gapdh* levels were similar in *Dfl* and wild type samples in the microarray data so this was used for normalisation. Triplicate PCR reactions using *Gapdh* were always run on the same plate as the test genes to allow an accurate comparison between samples.

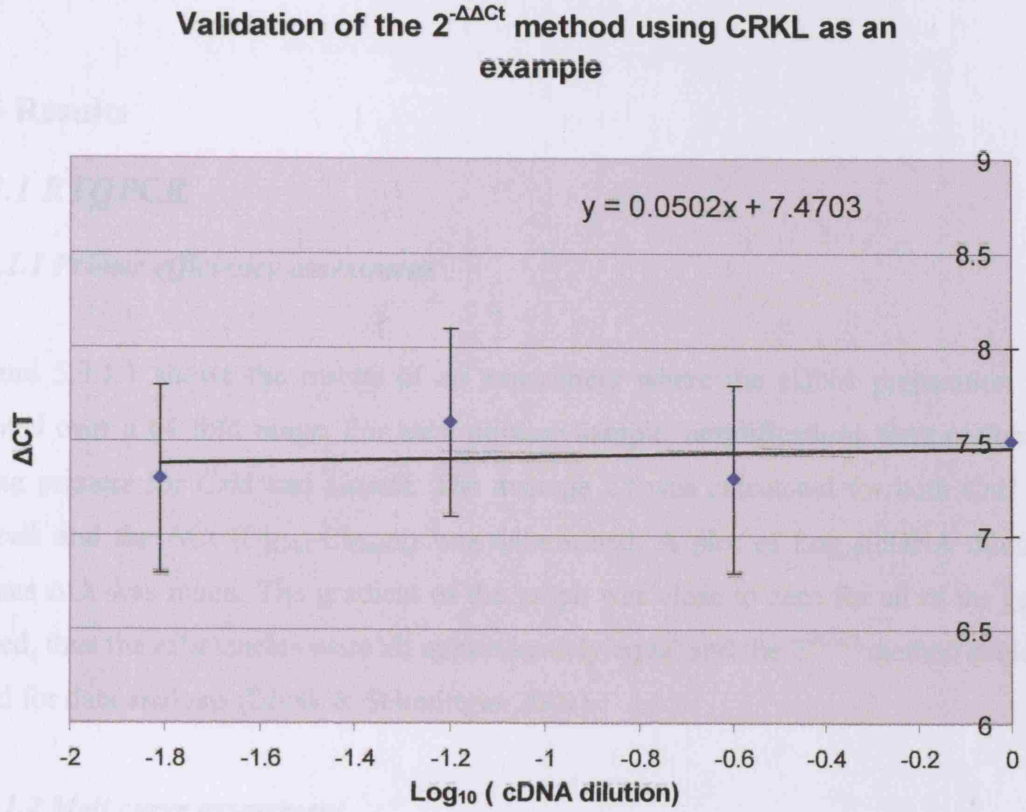
5.2.1.6 Primer efficiency assessment

For the $2^{-\Delta\Delta Ct}$ method to be valid, the amplification efficiencies of the target (gene under investigation e.g. *Crkl*) and control gene (*Gapdh* in this case) must be approximately equal. A sensitive method for assessing if two amplicons have the same efficiency is to look at how ΔCt varies with template dilution. If the absolute value of the slope is close to zero, the efficiencies of the target and reference genes are similar and the $2^{-\Delta\Delta Ct}$ calculation for the relative quantification of target may be used. This was done for all of the genes tested (figure 5.3.1.1).

5.2.2 In situ hybridisation

5.2.2.1 Probe design

Where possible clones for *in situ* hybridisation were requested and obtained from other laboratories. However if this was not possible the clones were ordered from the IMAGE consortium (http://www.geneservice.co.uk/home/). Probe design was important to minimise cross reactivity with other mRNA's. Sequences used were usually within the 3' UTR region of the mRNA. The sequence was always analysed by the NCBI BLAST facility to assess its specificity.



Raw Data

Dilution		Gapdh	Crkl	ΔCt
1x	Replicate 1	16.54873	23.55971	
	Replicate 2	16.26179	23.85555	
	Mean Ct	**16.40526**	**23.70763**	**7.3**
4x	Replicate 1	18.61265	26.04139	
	Replicate 2	18.47321	26.3318	
	Mean Ct	**18.54293**	**26.186595**	**7.6**
16x	Replicate 1	20.86688	28.25071	
	Replicate 2	20.98939	28.2633	
	Mean Ct	**20.928135**	**28.257005**	**7.3**
64x	Replicate 1	22.60917	29.97716	
	Replicate 2	22.94887	30.48441	
	Mean Ct	**22.77902**	**30.230785**	**7.5**

Figure 5.3.1.1

Validation of the $2^{-\Delta\Delta Ct}$ method: Amplification of cDNA diluted from 1-64 fold, using *Gapdh* and *Crkl* primers. Raw data is shown above. The Y axis shows the difference in Ct between *Gapdh* and *Crkl* and the X axis shows Log_{10}(dilution of the cDNA). The data were fit using the least squares linear regression analysis where the equation of the regression line is shown. As the gradient of the line is close to zero the efficiencies of the two genes are approximately equal and the $2^{-\Delta\Delta C}$t method of data analysis can be used.

5.3 Results

5.3.1 RTQPCR

5.3.1.1 Primer efficiency assessment

Figure 5.3.1.1 shows the results of an experiment where the cDNA preparation was diluted over a 64 fold range. For each dilution sample, amplifications were performed using primers for *Crkl* and *Gapdh*. The average Ct was calculated for both *Crkl* and *Gapdh* and the ΔCt (Ct_{Crkl}-Ct_{Gapdh}) was determined. A plot of Log_{10}(cDNA dilution) versus ΔCt was made. The gradient of the graph was close to zero for all of the genes tested, thus the efficiencies were all approximately equal and the $2^{-\Delta\Delta Ct}$ method could be used for data analysis (Livak & Schmittgen 2001).

5.3.1.2 Melt curve assessment

Two of the genes analysed had abnormal melt curve formation: *Pkrir* and *Ufd1l*. For *Pkrir* the primers were redesigned and the 2nd set of primers had a normal curve and could therefore be used. For *Ufd1l* the melt curve showed that there was a problem with likely primer-dimer formation and the formation of two PCR products. This was confirmed by running the PCR products on a gel. It is possible to change the conditions of the PCR reaction to eliminate the contribution the unwanted product makes to the fluorescent readings. This was done for *Ufd1l* by introducing an 8 second step at 79°C before the fluorescent read to melt the unwanted product. In addition halving the primer concentration of *Ufd1l* reduced the primer-dimers to a negligible amount. The melt curve analyses for *Ufd1l* and *Pkrir* before and after the changes described are shown in figure 5.3.1.2.

5.3.1.3 Example of the calculation of a relative expression level

cDNA amplified with primers for *DnaJb9* is used to describe the calculation involved in arriving at the relative change for a gene using the method devised by Livak *et al* (Livak & Schmittgen 2001). This example is illustrated in figure 5.3.1.3. Data where the standard deviation of the Ct values was greater than 0.3 were not analysed as they were considered too variable to be reliable.

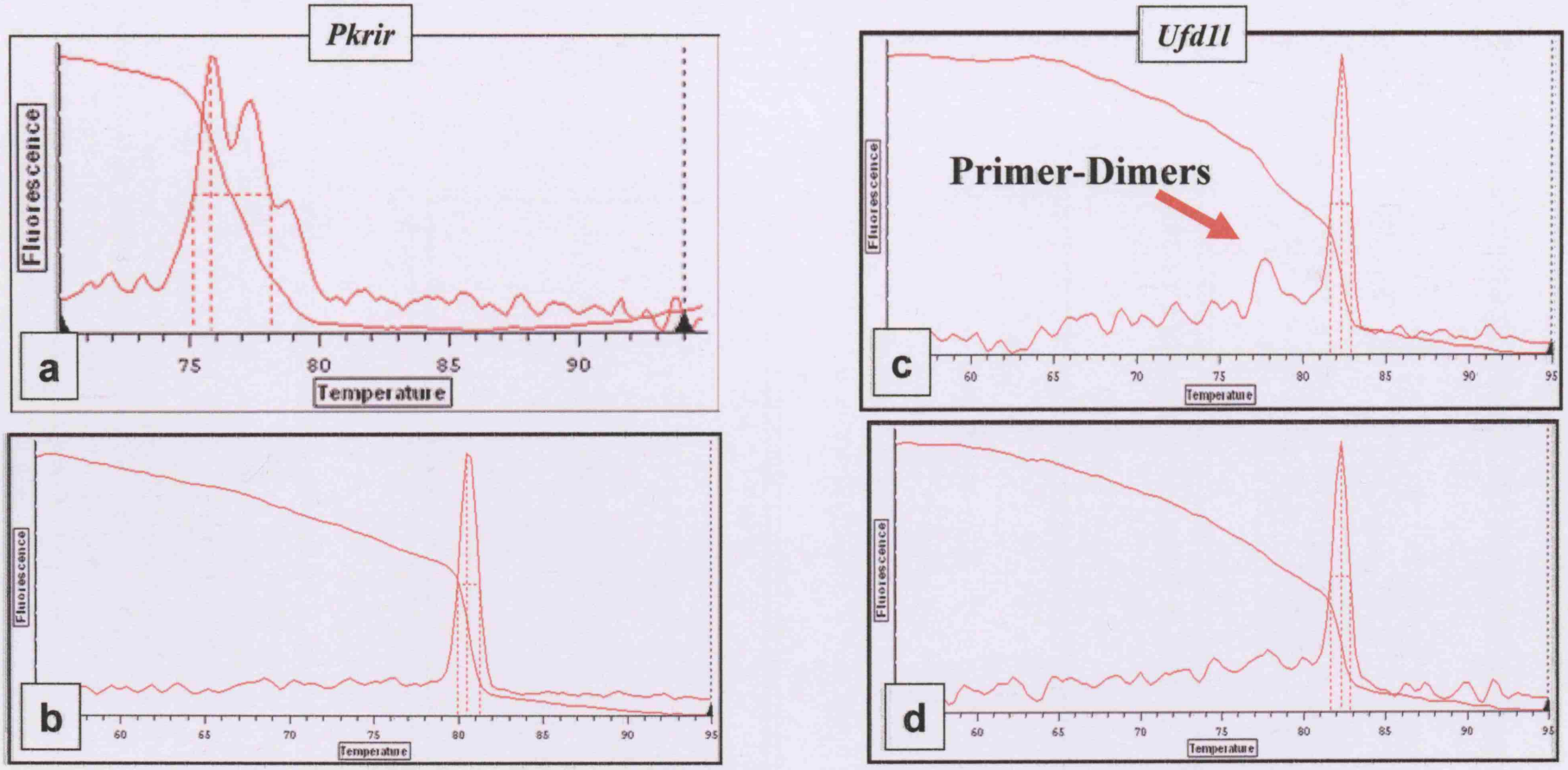


Figure 5.3.1.2

Melt curve assessments. a) *Pkrir* before primer redesign b) after primer redesign. c) *Ufd1l* before redesign d) after reducing the primer concentration and introducing an 8 sec melt step at 79°C read before the fluorescent read.

Figure 5.3.1.3Worked example of a relative change calculation.

	a		b		c		d
cDNA	**Ct value**	**cDNA**	**Ct value**	**cDNA**	**Ct value**	**cDNA**	**Ct value**
Wt Gapdh	23.88	Dfl Gapdh	24.27	Wt DnaJb9	29.89	Dfl DnaJb9	31.61
Wt Gapdh	23.64	Dfl Gapdh	24.34	Wt DnaJb9	30.18	Dfl DnaJb9	31.62
Wt Gapdh	23.91	Dfl Gapdh	24.34	Wt DnaJb9	29.99	Dfl DnaJb9	31.29
Mean Ct	**23.82**	**Mean Ct**	**24.32**	**Mean Ct**	**30.02**	**Mean Ct**	**31.50**
Ct SD	**0.15**	**Ct SD**	**0.04**	**Ct SD**	**0.15**	**Ct SD**	**0.19**

Relative amount of DnaJb9 in Dfl=$2^{-\Delta\Delta Ct}$

$$= 2^{-(a-b+d-c)}$$

Where:

a = mean Ct of Wt Gapdh =23.82

b = mean Ct of Dfl Gapdh =24.32

c = mean Ct of Wt DnaJb9 =30.02

d = mean Ct of Dfl DnaJb9 =31.50

$$= 2^{-(a-b+d-c)}$$

$$= 2^{-(23.82-24.32+31.5-30.02)}$$

$$= 2^{-0.98}$$

$$= 0.51$$

Figure 5.3.1.3

In this example the relative amount of the gene *DnaJb9* is calculated for the *Dfl* cDNA relative to the wild type (Wt) cDNA. The table shows the raw data downloaded from the Opticon RTQPCR machine. The Ct values are shown for the triplicate Wt and *Dfl* samples and the mean Ct and the standard deviations (SD) of the Ct values are shown. cDNA of an identical dilution was used for analysis by primers for *Gapdh* and *DnaJb9*. The mean Ct for each data set is used to calculate the relative amount of expression of *Dnajb9* in the *Dfl* sample compared to the wild type sample. This has been adjusted for the cDNA input amounts which is measured by the relative change in Ct values for the control gene *Gapdh*. In this example *DnaJb9* is expressed at 51% of wild type levels.

5.3.1.4 Results of genes followed up within the Df1 deleted region

The results of the genes followed up by RTQPCR from the *Df1* deleted region are detailed in the tables below. MA is the relative expression of the gene in the *Df1* sample compared to the wild type from the microarray (MA) result.

Affymetrix No.	Df1 deleted genes-statistically significantly downregulated	MA	RTQPCR
102403	cell division cycle 45 homolog (S. cerevisiae)-like (CDC45)	0.48	0.47
98573	RAN binding protein 1 (ranbp1)	0.54	0.24
160437	thioredoxin reductase 2 (TXNRD2)	0.58	0.29
98535	catechol-O-methyltransferase (COMT)	0.59	0.55

Table 5.3.1.4.1

The table above compares the MA relative expression (Df1/Wt) with RTQPCR relative expression (Df1/Wt) for the *Df1* deleted genes passing the Genespring statistical filtering applied.

Affymetrix No.	Df1 deleted genes-not passing statistical testing	MA	RTQPCR
Not on array	Tbx1	NA	0.49
93303	Ubiquitin dependent proteolytic protein (Ufd1l)	0.63	0.60
95025	DNA segment, Chr 16, human D22S680E, expressed (T10)	0.79	0.41
160574	Expressed sequence 2 embryonic lethal (Es2)	1.20	0.52

Table 5.3.1.4.2

The table above compares the MA relative expression (Df1/Wt) with RTQPCR relative expression (Df1/Wt) for the *Df1* deleted genes not passing the Genespring filtering applied.

These relative changes are summarised in figure 5.3.1.4 which also illustrates the relative distance of the genes to each other within the *Df1* deleted region.

5.3.1.5 Results of genes near to the Df1 deleted region

Affymetrix No.	Genes outside but near the Df1 deleted region	MA	RTQPCR
103848	v-crk sarcoma virus CT10 oncogene homolog (avian)-like (CRKL)	0.77	0.66
103009	histone cell cycle regulation defective homolog A (S. cerevisiae; HIRA)	1.02	1.05
93603	mitochondrial ribosomal protein L40 (Mrp140)/ NLVCF	0.95	1.1

Table 5.3.1.5

The table above compares the MA relative expression (Df1/Wt) with RTQPCR relative expression (Df1/Wt) for the genes outside but near the *Df1* deleted region.

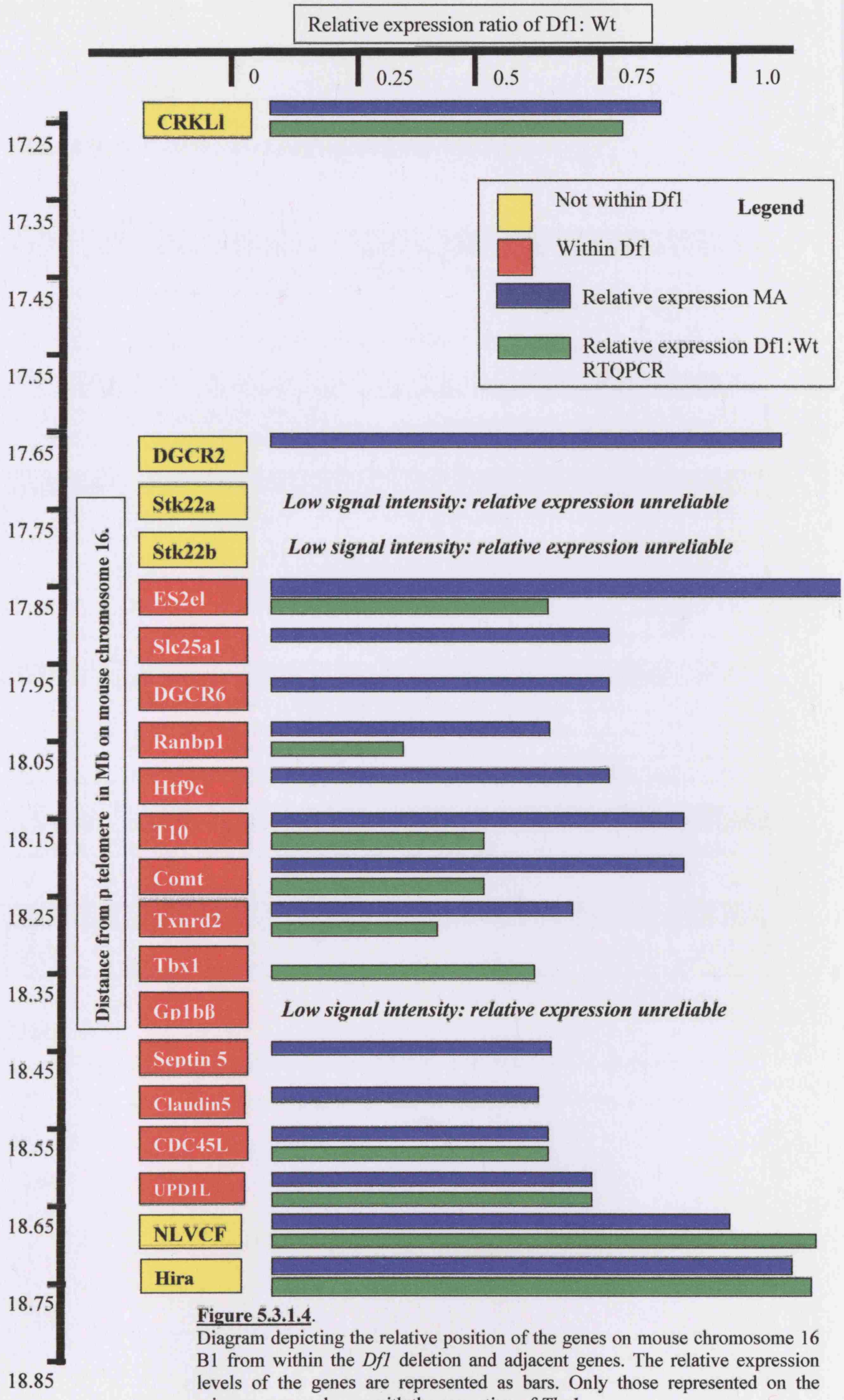


Figure 5.3.1.4.
Diagram depicting the relative position of the genes on mouse chromosome 16 B1 from within the *Df1* deletion and adjacent genes. The relative expression levels of the genes are represented as bars. Only those represented on the microarray are shown with the exception of *Tbx1*.

5.3.1.6 RTQPCR verification of differentially expressed genes

Downregulated Genes

Affy No.	Genes implicated in vasculogenesis/angiogenesis	MA	RTQPCR
96679	DnaJ (Hsp40) homolog, subfamily B, member 9 (DnaJb9/Mdg1)	0.46	0.7
104590	myocyte enhancer factor 2C (Mef2c)	0.58	1
100947	transcription factor 20 (Tcf20)	0.62	0.71
104635	gap junction membrane channel protein alpha 7-connexin 45 (Cx45)	0.65	0.66

Affy No.	Genes implicated in cardiogenesis	MA	RTQPCR
102419	retinoic acid receptor, gamma (rarg)	0.55	0.9
93669	SRY-box containing gene 11 (Sox11)	0.68	0.79

Affy No.	Transcription factors/cofactors	MA	RTQPCR
102085	insulinoma-associated 1 (IA-1)	0.29	0.65
96500	Gata5	0.36	0.72
93950	forkhead box A2 (Foxa2)	0.37	0.75
93701	SWI/SNF related, matrix associated, actin dependent regulator of chromatin (Smarca5)	0.59	0.67
103302	SRY-box containing gene 3 (Sox3)	0.6	0.58
161146	T-box 2 (Tbx2)	0.62	0.81

Affy No.	Protein Signalling pathway genes	MA	RTQPCR
97227	guanine nucleotide binding protein, alpha 12 (Gnbp12)	0.35	0.6
101831	Sonic hedgehog (Shh)	0.45	0.72
95680	protein phosphatase 1, regulatory (inhibitor) subunit 2 (PPIR2)	0.47	0.68
99975	protein-kinase, interferon-inducible double stranded RNA dependent inhibitor (PKRIR)	0.63	0.66

Affy No.	Apoptosis related/cell cycle related genes	MA	RTQPCR
104275	p53	0.7	0.77

Upregulated genes

Affy No.	Genes	MA	RTQPCR
93573	Mouse gene for Metallothionein-I (three exons).	2	0.67
99024	Max dimerization protein 4	1.3	1.14

Table 5.3.1.6

Comparison of the MA relative expression (Df1/Wt) with RTQPCR relative expression (Df1/Wt) for the potentially biologically relevant genes grouped by their biological pathway.

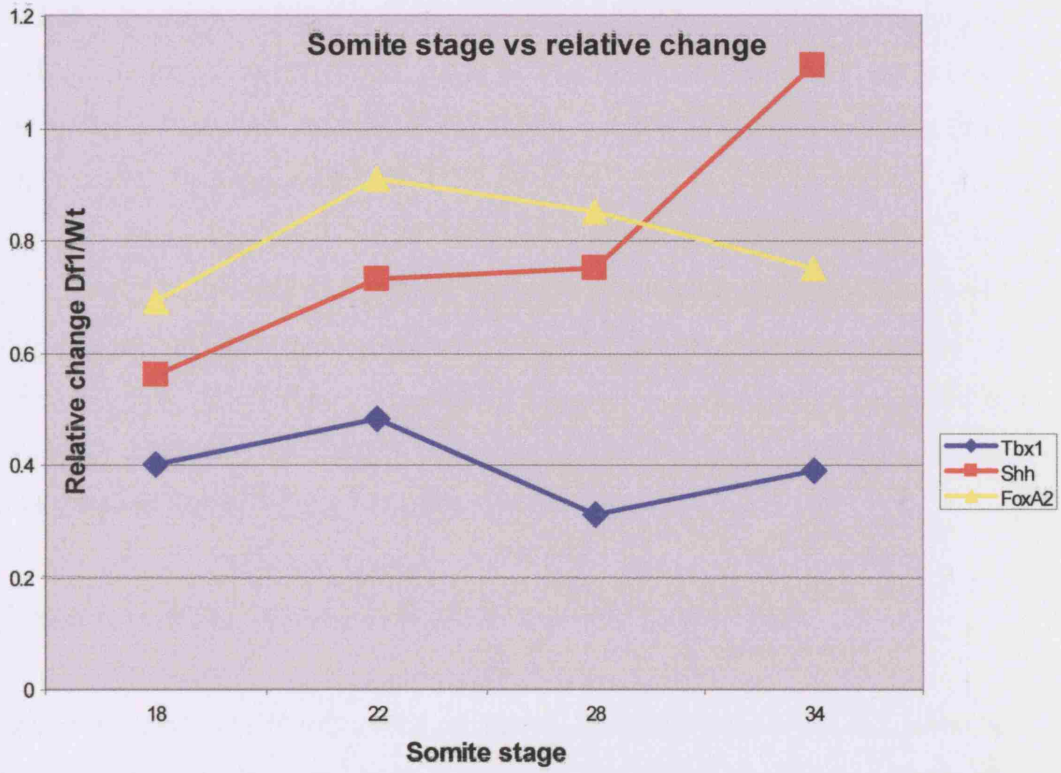
5.3.1.7 Results of target analysis in Tbx1 heterozygote E10.5 embryos.

		RTQPCR	
Affymetrix No.	Genes	Df1	Tbx1
96679	DnaJ (Hsp40) homolog, subfamily B, member 9 (DnaJb9/Mdg1)	0.7	1.03
104590	myocyte enhancer factor 2C (Mef2c)	1	0.65
100947	transcription factor 20 (Tcf20)	0.71	0.84
104635	gap junction membrane channel protein alpha 7-connexin 45 (Cx45)	0.66	0.89
102419	retinoic acid receptor, gamma (rarg)	0.9	0.92
93669	SRY-box containing gene 11 (Sox11)	0.79	0.57
102085	insulinoma-associated 1 (IA-1)	0.65	0.74
93950	forkhead box A2 (Foxa2)	0.75	0.56
93701	SWI/SNF related, matrix associated, actin dependent regulator of chromatin (Smarca5)	0.67	1.6
103302	SRY-box containing gene 3 (Sox3)	0.58	1.02
161146	T-box 2 (Tbx2)	0.81	0.56
97227	guanine nucleotide binding protein, alpha 12 (Gnbp12)	0.6	0.51
99975	protein-kinase, interferon-inducible double stranded RNA dependent inhibitor (PKRIR)	0.66	0.72
104275	p53	0.77	0.6
103848	Crkl	0.66	0.84

Table 5.3.1.7

Comparison of RTQPCR ratios (mutant/wild type) for cDNA from *Df1* embryos with cDNA from *Tbx1* heterozygote embryos.

The filtering applied to produce set 1a (downregulated genes) and 2 (upregulated genes) produced datasets of 141 (table 4.3.5.1.1) and 144 genes (table 4.3.5.1.2) respectively. Set 1b comprised of 58 genes, 3 of which were included in table 4.3.5.1.1. Genes from these lists were selected on the basis of known or potential biological function for further analysis by RTQPCR. 14 of 17 genes were confirmed as differentially expressed, with RTQPCR *Df1*:control ratios of <0.8 (table 5.3.1.6). These included *DnaJb9, Tcf20* and *Cx45*. RTQPCR failed to confirm downregulation of *Tbx2*, *Mef2c* and *RARγ*. The gene with the greatest fold change in set 1a was *IA-1* with a relative expression level of 0.29 in the *Df1* compared to wild type embryos. However, this degree of downregulation was not confirmed by RTQPCR, with an average relative change of 0.65. Thus, there were no genes represented on the U74Av2 array that convincingly demonstrated a non-linear response to the *Df1* hemizygosity.



Somite Stage	18	22	28	34
Tbx1	0.4	0.48	0.31	0.39
Shh	0.56	0.73	0.75	1.11
FoxA2	0.69	0.91	0.85	0.75

Figure 5.3.1.8

This graph shows the relative expression changes in *Tbx1, Shh* and *FoxA2* over a series of different stages from 18-34 somites. The table below shows the raw data from the RTQPCR experiment.

5.3.1.8 Results of investigation of FoxA2/Tbx1/ pathway

In analysing the list of 58 genes in set 1b, (derived by filtering for genes with expression levels <50% of wild type, but allowing detection p values >0.1) it was noticed that the gene *Foxa2* showed down regulation to 0.37 wild type levels (p=0.004) on the arrays, and 0.75 wild type levels on RTQPCR. As *Foxa2* had been proposed as a regulator of *Tbx1* investigation of whether a feedback loop might result in *Df1* embryos having *Tbx1* expression levels <0.5 was pursued.

In order to investigate the potential relative temporal relationship between *Shh, FoxA2* and *Tbx1* over time, an experiment was performed using cDNA from embryos across a somite series. Figure 5.3.1.8 shows the relative expression levels in *Df1* compared to wild type cDNA at the 18, 22, 28 and 34 somite stages. The graph shows an inverse relationship between *Shh* and *FoxA2* at the 34 somite stage (figure 5.3.1.8).

5.3.1.9 Comparison of the transcriptome of Df1 and wild type embryos

Cluster analysis of all genes expressed at above threshold levels (9224) failed to discriminate between wild type and mutant samples. Creating an experiment tree using either standard or Pearsons correlations did not cluster *Df1* and control arrays into two groups (Figure 5.3.1.9). Only when a set of 418 changed genes in *Df1* versus wild type (Cluster Set: see methods 2.3.6.4) were selected could the two genotypes separate into two groups, demonstrating the lack of major global changes between genotypes.

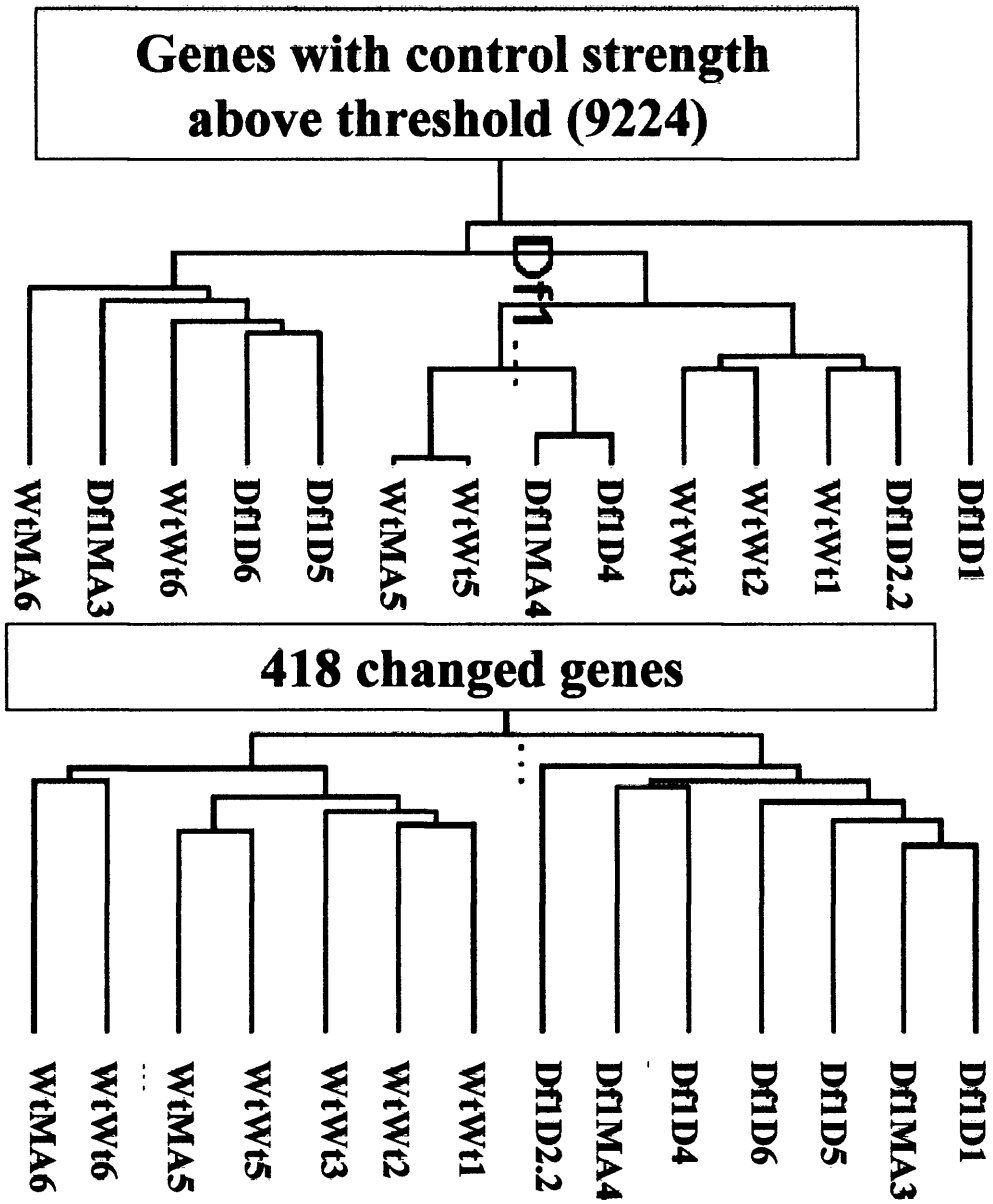


Figure 5.3.1.9

The upper panel shows an experimental tree clustered into Pearson's correlation. *Df1* and control arrays do not assort into two groups when the entire transcriptome of genes expressed with a control strength above threshold level are analysed. Only when the 418 most changed genes are selected as described in the text, and used to cluster the experiments, could the *Df1* arrays be reliably discriminated from wild type hybridisations (lower panel).

5.3.2 In situ hybridisations

Figure 5.3.2.1a

1) Connexin 45

Connexin 45 is expressed in the forebrain (fb); primitive atrium and primitive ventricle of the heart (h); forelimb (fl) and hindlimb (hl); 1st and 2nd branchial arches (I and II). There were no obvious differences noted on comparing the wild type to the *Df1* e10.5 embryos.

2) Rarγ

The probe for *Rarγ* showed strong expression in the forelimb and hindlimb, the anterior aspects of the 1st and 2nd branchial arches. No apparent differences were noted between the wild type and the *Df1* e10.5 embryos.

3) Foxa2

There was quite faint staining on the 1st and 2nd branchial arches. There was stronger staining in area "X" caudal to the branchial arches which represents the lateral plate mesoderm. The stripe down the midline axis of the embryo represents *FoxA2* staining in the notochord (NC). No significant differences were noted between the wild type and *Df1* embryos.

4) Sense control

No significant staining was noted on the sense control. Dye trapping was noted in the otic vesicle (ov).

Figure 5.3.2.1b

1) Tcf20

Staining with the *Tcf20* antisense probe showed expression in the somites (S), the forelimb (fl) and hindlimb (hl) buds, the 1st and 2nd branchial arches (I and II) and the outflow tract of the heart (OT). No apparent differences were noted between the wild type and the *Df1* e10.5 embryos.

2) Tbx2

Expression of this riboprobe was strong in the maxillary {I (Max)}, and mandibular {I (Mand)} prominence of the 1st branchial arch; anterior aspect of the 2nd branchial arch; rostral and caudal margins of the forelimb bud and the hindlimb bud. No apparent differences were noted between the wild type and the *Df1* e10.5 embryos.

3) Sox3

Strong expression was seen in the forebrain (fb), the developing neural tube (nt) and the trigeminal ganglia (tg) supplying the 1st arch. There was apparent expression in the otic vesicle (ov), and hindbrain (hb), but this could be artefact due to dye trapping. There were no obvious significant differences between wild type and *Df1* samples. The apparent decreased staining in the *Df1* embryo, was not seen in all embryos and is probably a reflection of the slight increase in stage of this embryo.

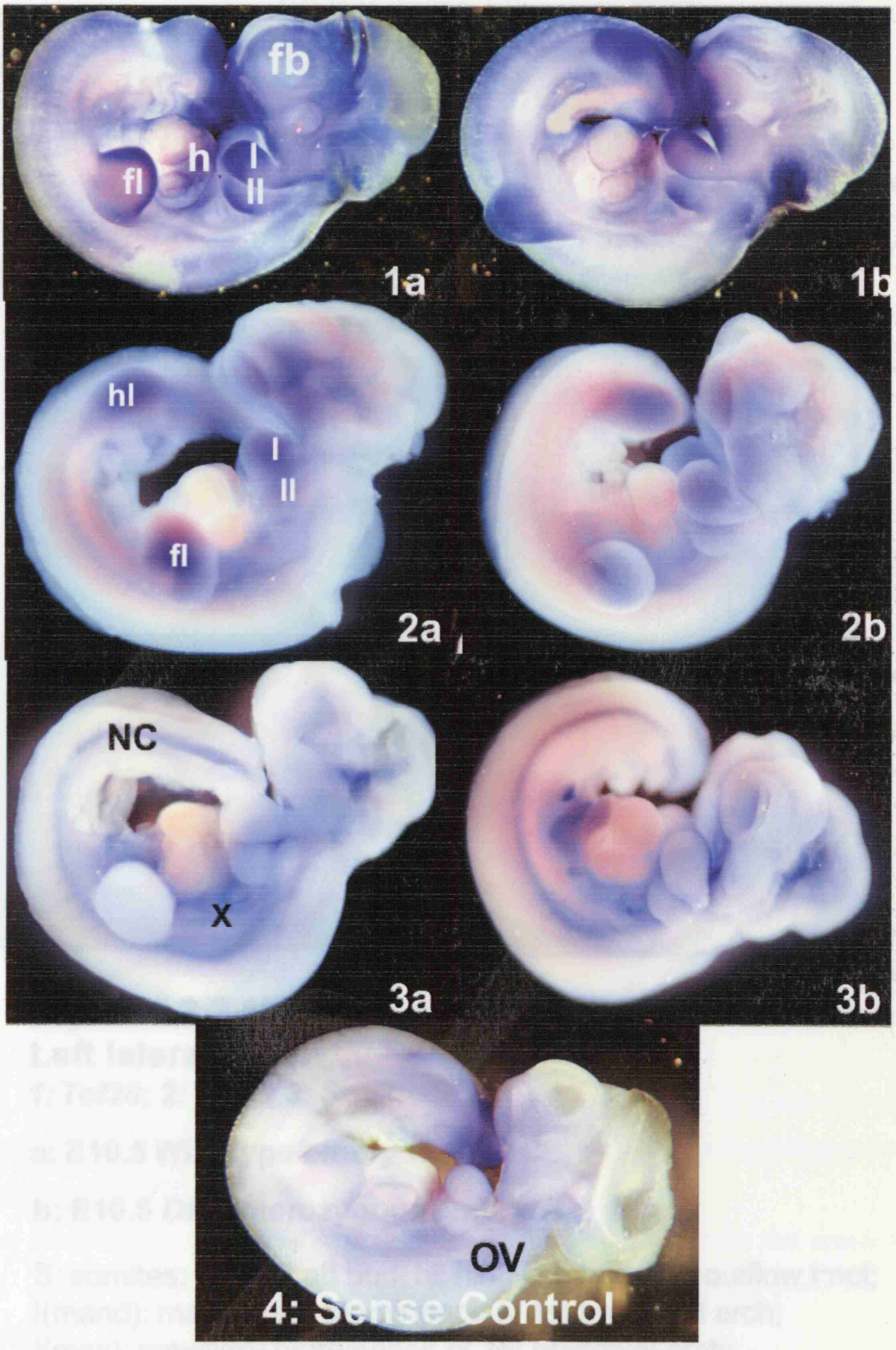


Figure 5.3.2.1a

Whole mount in situ hybridisations. Left lateral views.

1: Connexin 45; 2: Rarg; 3: FoxA2; 4: Sense Control

a: E10.5 Wild Type embryos

b: E10.5 *Df1* heterozygous embryos

fl: forelimb; hl: hindlimb; fb: forebrain; h: heart;
I: 1st branchial arch; II: 2nd branchial arch;
NC: notochord; OV; otic vesicle: X: lateral plate mesoderm

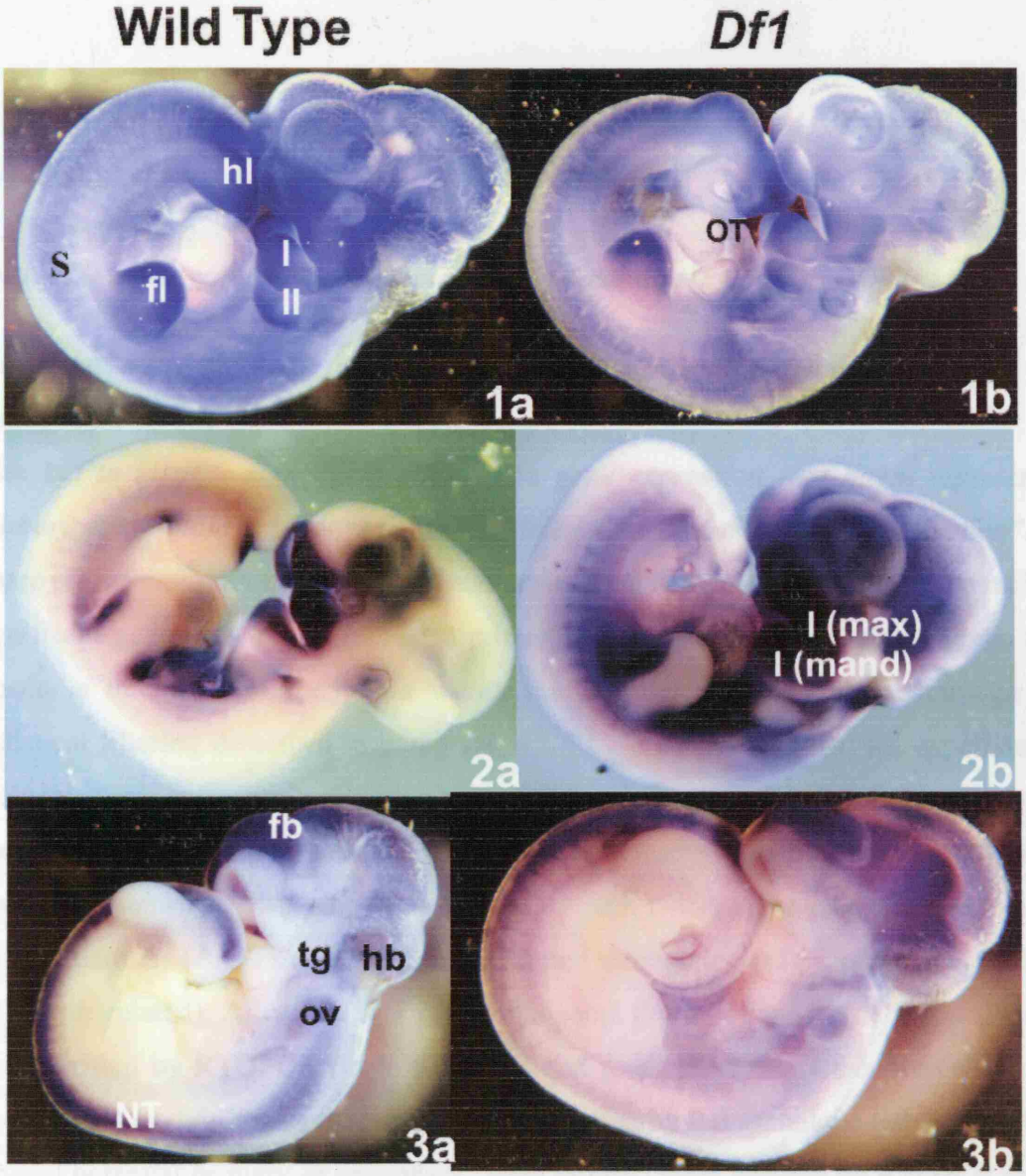


Figure 5.3.2.1b Whole mount in situ hybridisations. Left lateral view.
1: Tcf20; 2: Tbx2; 3: Sox3.

a: E10.5 Wild type embryos,

b: E10.5 *Df1* heterozygous embryos.

S: somites; fl: forelimb bud; hl: hindlimb bud; OT: outflow tract;
I(mand): mandibular prominence of 1st branchial arch;
I(max): maxillary prominence of 1st branchial arch;
II: 2nd branchial arch; S: somites; NT: neural tube;
fb: forebrain; ov: otic vesicle; hb: hindbrain; tg: trigeminal ganglia

5.4 Discussion

5.4.1 RTQPCR

5.4.1.1 Have the results answered the experimental aims?

The three main aims of this experiment are described in paragraph 4.1.3.3 and summarised below.

Aim 1a) ***to identify significant changes in gene expression that might represent a primary or secondary response to the*** *Df1* ***deletion-so called "target identification".***

There are several reasons why the up and down regulated gene sets are likely to include significant genes that might represent primary or secondary *Df1* deletion targets. One reason is that 69% of the genes within the *Df1* deleted region came through the statistical filtering described in paragraph 4.3.5.1. This indicated that the experimental protocols were sufficiently sensitive to detect two-fold expression level changes. Other reasons are that known targets such as *FoxA2* came through the filtering in set1b; some genes represented by two probe sets filtered through twice; and a pair of genes that are binding partners (*PKRIR* and *PKIR*) have filtered through in set 1a (upregulated) and set 2 (downregulated). All of the *Df1* genes that came through the MA statistical filtering were verified by RTQPCR (table 5.3.1.4.1) and four others that either weren't on the MAs (*Tbx1*) or didn't pass the filtering were verified by this technique (table 5.3.1.4.2). Several potential primary or secondary "gene targets" were verified by RTQPCR (table 5.3.1.6) and are discussed below (paragraph 5.1.4.3).

Aim 1b) ***to identify whether hemizygosity for*** *Df1* ***genes results in a significantly greater than two-fold change in the expression level of any other gene represented on the array: "amplification of haploinsufficiency".***

There are four genes in table 4.3.5.1.1 (downregulted in *Df1* embryos) that appear to have a relative expression level of below 0.4 in *Df1* embryos. These genes are as follows:

Gene	Affy no.	Relative expression *Df1*: Wt
Insulinoma-associated 1 (IA-1)	102085	0.29
Guanine nucleotide binding protein, alpha 12 (Gnbp12)	97227	0.35
GATA-5	96500	0.36
FoxA2	93950	0.37

Although these genes appear to represent "amplification of haploinsufficiency", RTQPCR did not confirm this level of decreased expression in *Df1* cDNA compared to wild type (table 5.3.1.6). In addition *Gnbp12* is represented on the array by two probe sets and the second probe set records a relative expression reading of 0.71. This suggests that actual the relative change ratio should only be used as a guide, with *significant* changes offering more accuracy than the *degree of* change.

Aim 2) ***to determine whether the Df1 deletion affected the expression of genes adjacent to, but not within the deletion.***

The analysis of patients with suspected 22q11DS has revealed a series of atypical deletions and apparently balanced translocations in patients with differing phenotypes related to 22q11DS (figure 1.4.2.2). Some of these rearrangements do not directly disrupt *TBX1*, suggesting that the disturbance of chromosome architecture within 22q11 might affect the expression of genes at some distance from the rearrangement. To analyse the effect of the *Df1* deletion on neighbouring loci the expression level of the *Mrp140* (*Nlvcf*) gene, the closest gene to the distal endpoint, some 40kb from *Ufd1l,* was analysed. In addition the expression level of the *Hira* gene was analysed as it is required for normal embryogenesis and has been studied as a potential contributor to the human syndrome. *Hira* maps 50kb distal to the *Df1* deletion, adjacent to *Mrp140.* Proximal to the breakpoint *Stk22a* and *Stk22b* were examined, both within 20kb of *Es2. Stk22a* and *Stk22b* show relative expression levels of 1.73 and 1.79 respectively, but these are misleading as the signal strength for these genes was below the threshold for reliability. As shown in table 4.3.5.2c, none of the genes immediately adjacent to the deletion and expressed at E10.5 had altered expression levels, adding evidence that the phenotype of *Df1* is solely attributable to the engineered deletion itself.

Crkl is a gene located approximately 400kb to the *Df1* deletion and it has been proposed that deletion of *CRKL* may contribute to the human 22q11DS (Guris et al. 2001). Microarray analysis gave a differential expression level of 0.77 and this was corroborated by RTQPCR (0.66) shown in table 5.3.1.5.

Crkl-/- mutants have interrupted aortic arches, absent parathyroid glands, hypoplastic or ectopic thymus glands and craniofacial malformations (Guris et al. 2001). The distance of *Crkl* to the *Df1* deletion makes it unlikely to be affected by position effect of the *Df1*

mutation given the negative data for genes immediately adjacent to *Df1*. This raises the possibility that *Df1* affects *Crkl* expression. In support of this, it has recently been reported that compound heterozygosity of *Tbx1* and *Crkl* has been shown to cause abnormal patterns of the great vessels such as interrupted aortic arch type B, abnormal origin of the right subclavian artery together with a hypoplastic or aplastic thymus, defects more severe than those seen in *Df1/+* or *Tbx1/+* embryos (Guris et al. 2006).

Aim 3) ***to establish whether genes within the Df1 deletion that are not differentially expressed on the array exhibit dosage compensation or upregulation of the non-deleted allele.***

Dosage compensation is well studied in organisms such as Drosophila, where regulatory mechanisms result in hypertranscription from the single male X to give the equivalent expression of X-chromosome genes in both sexes (Chiang & Kurnit 2003). In mice dosage compensation of single autosomal genes has been demonstrated in targeted mutations of the *Brn3a* transcription factor (Trieu et al. 2003).

Of all the genes shown in table 4.3.5.2a and b, three genes appeared to possibly exhibit the phenomenon of upregulation of the non-deleted allele or dosage compensation: *UfD1L, T10* and *Es2*. However RTQPCR for these three genes confirmed that they do have decreased expression levels in *Df1* embryos (figure 5.3.1.4.2). This confirms the hypothesis that the probe design on the arrays for *UfD1L* and *T10* is one reason for them not showing reduced expression in the microarray experiment. In the case of *Es2* probe design was good (table 4.3.5.2c), but *Es2* is at the centromeric boundary of the *Df1* deletion. It can be hypothesised that ectopic expression of the remaining 3' end of *Es2* led to hybridisation to the array probes and the near wild type expression ratio of *Es2* in *Df1* embryos. This hypothesis is supported by data from Ivins *et al* who also found a lack of *Es2* downregulation on two array platforms (Ivins et al. 2005). RTQPCR analysis using probes designed to the 3' UTR gave expression levels of 0.52 in *Df1* embryos. Thus there was no evidence for dosage compensation for any of the *Df1* genes analysed.

5.4.1.2 Comparison of the transcriptome of Df1 and wild type embryos

Global comparisons of transcriptome profiles did not discriminate between the *Df1* and control groups of arrays. This is in contrast to the effect of *Ts65Dn* on the cerebellar genome of the mouse, the published study with greatest similarity to our own (Saran et al. 2003). In this work, a similar number of probes were analysed and selection of 1532 discriminating probes (those most consistently changed between genotypes) grouped the euploid and trisomy targets into two discrete clusters. In this experiment, it was found that a selected set of just over 400 changed probes was required to distinguish the two groups (figure 5.3.1.9). Removal of the most changed 1532 genes from the cluster analysis of cerebellar transcriptomes did not abolish the clustering of trisomic from euploid targets, which further emphasizes the more subtle changes caused by the *Df1* deletion versus trisomy 21. This is most readily explained by considering the differences in the number of genes in genomic dosage imbalance (124 in *Ts65Dn* mice versus 22 in *Df1*) and reflects the subtlety of the phenotype in *Df1* mutants.

A similarity with the work on the *Ts65Dn* model was the relative lack of genes with a greater than two fold variation. The set 1a and 2 filtering and statistical analysis employed here identified 141 and 144 genes expressed +/- 25% wild type levels respectively. One of my aims was to test the hypothesis that the *Tbx1* haploinsufficiency, in the context of concomitant deletion of adjacent genes, results in an alteration of expression greater than 50% in a critical target gene or genes (the amplification hypothesis). This phenomenon has been noted for a related gene, *Tbx5*, mutations of which are associated with the heart defects in Holt-Oram syndrome (Li et al. 1997b). Mice carrying a heterozygous loss of function mutation of the *Tbx5* show a large decrease in cardiac atrial natriuretic factor (ANF) expression and the binding of Tbx5 to recognition sites within the ANF promoter was a non linear function of Tbx5 concentration (Bruneau et al. 2001). *Connexin 40* expression was also substantially lower in this model (Bruneau et al. 2001). Just two genes on the array were above threshold levels for robust detection of expression changes and had expression ratio <0.3 (*IA-1:insulinoma associated-1*) or >2 (*GM2; ganglioside activator protein 2*), and neither seems a likely major player in the observed phenotype.

5.4.1.3 Differentially expressed genes: downregulated

The lists of other significantly changed genes were examined for their potential biological role and studied further with RTQPCR (table 5.3.1.6). No single biochemical pathway could be discerned as being particularly affected by the aneusomy, although three validated genes discussed below are implicated in angiogenesis. Given the existence of good positive controls for downregulated genes (the *Df1* deleted genes) and that any direct targets of *Tbx1* represented within the dataset would likely be downregulated {as *Tbx1* is known to be a transcriptional activator {Ataliotis et al. 2005}}, validation was biased towards potentially down regulated genes.

5.4.1.3.1 Genes implicated in vasculogenesis/angiogenesis

Connexin 45 (*Cx45*) and *Dnajb9/Mdg1* were found to be downregulated by microarray and RTQPCR, which might reflect a primary or secondary disturbance of pathways involved in vascular development. The *Dnajb9* gene encodes a protein referred to as microvascular differentiation gene 1, as it was isolated during a screen for genes induced during an *in vitro* model of angiogenesis (Prols et al. 2001). As *Dnajb9* is expressed in endothelium, and the endothelium of the 4th PAA appears to form normally in *Df1* embryos, the downregulation observed is likely to be a secondary event.

At mouse E10.5 *Cx45* is strongly expressed in the 1st and 2nd branchial arches, the forebrain, the heart, the limb buds and the vascular smooth muscle cells the dorsal aortae. Homozygous mutants die between E9.5 and E10.5 with severe vascular defects in the placenta and embryo (Kruger et al. 2000), together with endocardial cushion defects of the heart (Kumai et al. 2000). Although endothelium forms normally, the smooth muscular layer surrounding the major arteries fails to develop. The decrease in expression level of *Cx45* detected could be due to a reduction in smooth muscle cells in the *Df1* embryos surrounding the arch arteries. However this subtle difference is unlikely to be detectable in branchial arch dissections due to the abundance of other tissue types.

Null *Cx45* mutants also demonstrated heart defects with variable incidence of outflow tract dilatation, narrowing of the dorsal aortae and narrowing of the connection of the

PAAs with the dorsal aorta. In addition the connection of the branchial arch arteries with the dorsal aortae was narrow or absent in some *Cx45* null embryos. This indicates that *Cx45* exhibits a dose dependent regulation of blood vessel formation (Kruger et al. 2000) and that a moderate reduction in expression such as that observed in this MA and RTQPCR data, may in conjunction with other changes be of physiological significance. Interestingly, *Dfl* mutants show a defect in the differentiation of vascular smooth muscle cells in the 4th PAA. Thus, down regulation *of Cx45* could contribute to the phenotype observed in *Dfl*/+ embryos.

Tcf20, a transcription factor downregulated in *Dfl* embryos, may regulate tissue remodelling through its action at the stromelysin-1/matrix metalloproteinase (MMP) 3 promoters (Rekdal et al. 2000). Interestingly Cai *et al* have demonstrated that a synthetic MMP inhibitor decreased early neural crest migration in chick embryos (Cai & Brauer 2002). It is therefore possible that in *Dfl* mice, *Tcf20* is inhibited resulting in inhibition of *stromelysin-1* expression with a resultant effect on neural crest cell migration.

Mef2c, although an exciting candidate gene given its role in the anterior heart field formation (Dodou et al. 2004), did not verify with RTQPCR.

5.4.1.3.2 Genes implicated in cardiogenesis

Sox11 is another downregulated gene, which potentially contributes to the Dfl phenotype. *Sox11* is expressed at many sites undergoing tissue remodelling during embryogenesis. Homozygous targeted mutants are cyanotic and die at birth, having developed abnormalities in the outflow tract and ventricular septal defects (Sock et al. 2004).

Rarγ is a key player in cardiogenesis (Romeih et al. 2003), but its decreased expression was not confirmed by RTQPCR (relative expression 0.9) or *in situ* hybridisation (figure 5.3.2.1a).

5.4.1.3.3 Transcription factors/co-factors

GATA5 is a transcription factor that physically interacts with cardiac T-box transcription factor, *TBX20* (Stennard et al. 2003) to synergistically activate gene expression. *GATA5* has specifically been implicated in endocardial differentiation (Afouda et al. 2005). Its downregulation in the *Df1* embryos may account in a direct or indirect way for the cardiac phenotype seen.

Sox3 is a transcription factor expressed in the epibranchial placodes which give rise to the sensory neurons of the distal cranial ganglia (Abu-Elmagd et al. 2001). From the facial asymmetry that some 22q11DS patients have, and the misdirected migration of cranial nerve ganglia in *Tbx1* null mice (Vitelli et al. 2002b), it was considered a good candidate for follow up. Two different probes for *Sox3* detection came through the statistical analysis as being consistently downregulated in *Df1* embryos. One probe was also downregulated in *Df1/Tbx1* compound heterozygous embryos examined at E9.5 by Dr Ivins (Ivins et al. 2005). This apparent downregulation was not seen in *Tbx1* nulls (Ivins et al. 2005) or *Tbx1* heterozygous embryos (table 5.3.1.7) suggesting an effect of other genes within *Df1* besides *Tbx1*. Although *Sox3*'s potential downregulation was confirmed by RTQPCR, no significant differences were noticed on whole mount *in situ* hybridisation (5.3.2.1b). Furthermore a mouse knock out of *Sox3* in *Tbx1* expressing cells using the cre/floxed system did not result in a phenotype (Robin Lovell-Badge personal communication).

Tbx2 was considered a good candidate for follow up as it is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract (Harrelson et al. 2004). Its apparent downregulation was not confirmed by RTQPCR or *in situ* hybridisation (figure 5.3.2.1b). Affymetrix probe inspection revealed specificity to *Tbx2* with no crossover recognition of *Tbx1*.

5.4.1.4 Interpretation of FoxA2/Shh/Tbx1 results

Tbx1 is regulated by *Sonic Hedgehog (Shh)* during pharyngeal arch development (Garg et al. 2001). *FoxA2* has an intermediary role in *Shh* regulation of *Tbx1* (Yamagishi et al. 2003b). As well as *FoxA2* being upstream of *Tbx1* it has also recently been shown to be

a *Tbx1* target in *Tbx1* hypomorphs that have *Tbx1* expression levels at 25% of wild type (Hu et al. 2004). It is therefore postulated that *FoxA2* and *Tbx1* are involved in a reinforcing autoregulatory loop (Hu et al. 2004).

Shh and *FoxA2* had both come through in set 1b as statistically significant targets. In order to investigate their temporal relationship with *Tbx1*, a RTQPCR experiment with RNA from somite stages 18, 22, 28 and 34 was designed. For each stage relative expression of *Shh*, *FoxA2* and *Tbx1* was examined (figure 5.3.1.8). In interpreting this data it should be born in mind that inspection of such subtle changes in expression are probably beyond the limits of the sensitivity of RTQPCR, although general changes and trend interpretation is probably possible. With this in mind it is noted that as *Shh* expression increases towards the later somite stages, *FoxA2* expression diminishes (figure 5.3.1.8). Generally the reduced expression of *Tbx1* below 50% as *FoxA2* falls could also be a consequence of the autoregulatory loop.

5.4.1.5 Comparison between Df1 and Tbx1 expression profiles

As the *Tbx1* null mouse model recapitulates all of the pharyngeal, cardiac and craniofacial features of the *Df1* mouse, it is useful to compare the expression profiles of the two mouse models. A similar microarray experiment to my own has been conducted by Dr Sarah Ivins at MMU (Ivins et al. 2005). Comparison of downregulated genes that were in common with my dataset only revealed *Sox3*. RTQPCR analysis of Dr Ivins' potential *Tbx1* target genes in my *Df1* embryos did not show any significant downregulation (data not shown). This is perhaps even more surprising given that Dr Ivins used *Df1/Tbx1* transheterozygotes in order to utilise the haploinsufficient *Df1* genes in statistical filtering.

There are however, obvious differences between the two mouse models. *Tbx1* null embryos have 2nd pharyngeal arch hypoplasia and absence of the 3rd and 4th PAAs. Branchial arch dissections were also done differently (Ivins et al. 2005) and involved a more refined region of the embryo not encompassing the neural tube. Due to hypo/aplastic pharyngeal arteries Dr Ivins ran into the problem of tissue loss when comparing expression in wild type embryos with the *Tbx1* nulls.

In order to address the question of whether there were any targets in common between the *Df1* and *Tbx1* mice, similar branchial arch dissections on E10.5 *Tbx1* heterozygote mice to the ones performed on the *Df1* embryos were conducted (figure 2.3.4.3). Expression of the downregulated target genes in the cDNA of the *Tbx1* heterozygote embryos was analysed by RTQPCR. The results of this analysis are shown in table 5.3.1.7. Arbitrarily using the threshold of <0.8 as significant, 54% (7/13) of the RTQPCR confirmed *Df1* target genes also showed reduced expression in the *Tbx1* heterozygous embryos. It is possible that the other six genes represent direct or indirect targets of the *Tbx1* adjacent genes within the *Df1* deletion. A microarray study performed by Dr Ivins' using *Tbx1* heterozygous embryos compared with wild types did not reveal many significantly changed genes (Dr Ivins: work unpublished). This could be explained by the fact that haploinsufficiency of *Tbx1* results in a less severe phenotype than the *Df1* mice, with 25% of the *Tbx1* heterozygote mice having aortic arch patterning defects, but none having craniofacial or glandular defects (Merscher et al. 2001);(Kochilas et al. 2002);(Lindsay et al. 2001).

5.4.1.6 Differentially expressed genes: upregulated

An approximately equal number of upregulated and downregulated genes (n=140) came through the statistical filtering stage. Although follow up was bias towards downregulated genes, validation was attempted for a couple of upregulated genes: *Mad4* and *MetI*, but their upregulation was not confirmed by RTQPCR (table 5.3.1.6). Inspection of the raw microarray upregulated gene list (table 4.3.5.1.2) reveals a number of genes that are released in response to cellular stress. Six acute phase response genes and five other stress response genes (including *Metallothionein I and II*) were statistically increased in the microarray experiment. Examples of genes upregulated in the *Df1* embryos include the haemoglobin genes, insulin like growth factors and serum deprivation genes. Also quiescence inducing genes such as *Quiescin Q6* and *Decorin* were increased relative to the wild type embryos. This could be a reflection of the hypo/aplastic 4th pharyngeal arch arteries present in the *Df1* embryos as the embryos tries to protect against the tissue hypoxia this would produce. Metallothioneins act as antioxidants at times of cardiac stress in diabetic mice (Ye et al. 2005) and it is possible the same mechanism was occurring in the *Df1* embryos. The hypoxia in these embryos could have caused the surrounding cells to maximise their ability to survive by producing growth factors and stopping entry into mitosis.

Caveolin is an intriguing upregulated gene as it has been shown to produce endothelial proliferation and differentiation and its expression has been shown to enhance endothelial capillary tubule formation (Liu et al. 2002). It is possible that in the stressed state of the embryos with small pharyngeal arch arteries there is upregulation of this gene to initiate other methods of tissue oxygenation via increased capillary formation.

5.4.2 In situ hybridisation results

Generally no significant differences were observed between whole mount *in situ* hybridisations of *Df1* E10.5 embryos compared with wild type controls (figure 5.3.2.1a and b). This was anticipated given the subtle changes in expression that are expected. Sectioning of the embryos might have shown more subtle changes, but unfortunately time constraints did not allow this.

5.4.3 Conclusions and future work

It has been demonstrated that the fourth pharyngeal arch artery hypo/aplasia observed in *Df1* heterozygotes is fully penetrant at E10.5 but recovers as gestation proceeds such that approximately 20% of newborn mice have any evident cardiovascular defect. Presumably, the recovery is preceded by compensatory changes in gene expression in relevant tissues, and some of the changes in gene expression observed here might reflect this phenomenon. Thus, it will be interesting to examine whether there is any epistatic interaction between the differentially regulated genes detected here and the *Df1* or *Tbx1* mutation. For instance, if the downregulation of *Cx45* is a compensatory response, *Cx45/Tbx1* transheterozygotes would be expected to have a milder phenotype. On the other hand if gene such as *Cx45* is a direct target of *Tbx1* then transheterozygotes may have a more severe phenotype.

This work demonstrates that two fold expression level changes can be reliably detected in E10.5 *Df1* embryos, and that there is absence of dosage compensation. Overall, the transcriptome of *Df1* embryos was not grossly altered and few genes had a significantly changed expression level. Given the subtlety of the phenotype observed in *Df1* heterozygote embryos this result is perhaps not unexpected. While it is possible that

some of the RTQPCR validated genes, and others in sets 1a, 1b and 2, represent direct transcriptional targets of *Df1* deleted genes including *Tbx1*, additional biochemical approaches will be required to address this question. As *Tbx1* target genes may also be expressed in non-*Tbx1* expressing tissues at E10.5 the differential regulation of such targets will be diluted using the methods employed here, despite the dissection of the pharyngeal apparatus containing regions. These genes may well be downregulated in the dataset, but fail to meet the filtering and statistical cut offs applied to reduce noise to acceptable levels. It is clear that additional approaches will be required to identify direct targets of *Df1* deleted genes. Currently work is ongoing at MMU to examine *Tbx1* null embryos and cell lines with inducible expression of *Tbx1*. Differentially regulated genes obtained in these independent experiments will be compared with the wider sets of potentially differentially regulated genes identified here with a view to prioritising common genes for further analysis.

Chapter 6 Identification of *Pax3* target genes

6.1 Introduction

Pax3 is a paired-box and homeodomain containing transcription factor expressed in the region of the dorsal neural tube that gives rise to migrating neural crest populations (Goulding et al. 1991). Mutations in *Pax3* result in neural crest defects in mouse and man. *Splotch* mice result from a spontaneous heterozygous mutation in *Pax3* and are recognised by a white belly spot (Figure 6.1a). Homozygous *Splotch* embryos surviving to E13/14 exhibit conotruncal heart defects and abnormalities of the thymus, thyroid and parathyroid glands, similar to those seen in the 22q11DS (Franz 1989).

6.1.1 Relevance of experimental design to 22q11 Deletion Syndrome

As the *Pax3* homozygous embryos provide another mouse model for 22q11DS and as there was a delay in the arrival of the *Df1* mouse colony from Baylor due to flooding (Schub 2002), the decision was made to embark on a microarray study of *Pax3* E10.5 embryos to look for possible targets that may have relevance to the *Df1* microarray study (at the time it was not known that the defects seen in the *Df1* mouse were an indirect effect on neural crest cells). In addition there have been a number of publications of possible *Pax3* downstream targets in the literature, providing an opportunity to test the experimental sensitivity of the microarrays. The *Pax3* mouse colony was in house and the study was done in collaboration with Dr D. Henderson from Prof. A. Copp's laboratory at the Institute of Child Health.

6.1.2 The Splotch mouse

Six classical and radiation induced alleles of *Splotch* have been described: *Sp, Sp-delayed, Sp-retarded,* Sp^{1H}, Sp^{2H} and Sp^{4H}, which all have mutation in the *Pax3* transcription factor (Chalepakis et al. 1993). Sp^{2H} is a radiation induced allele at the *Pax3* locus that arose on the C3H/101 background (Beechey 1986).

Figure 6.1 Selection of figures demonstrating the abnormalities seen in the *Pax3* deficient mouse and human.

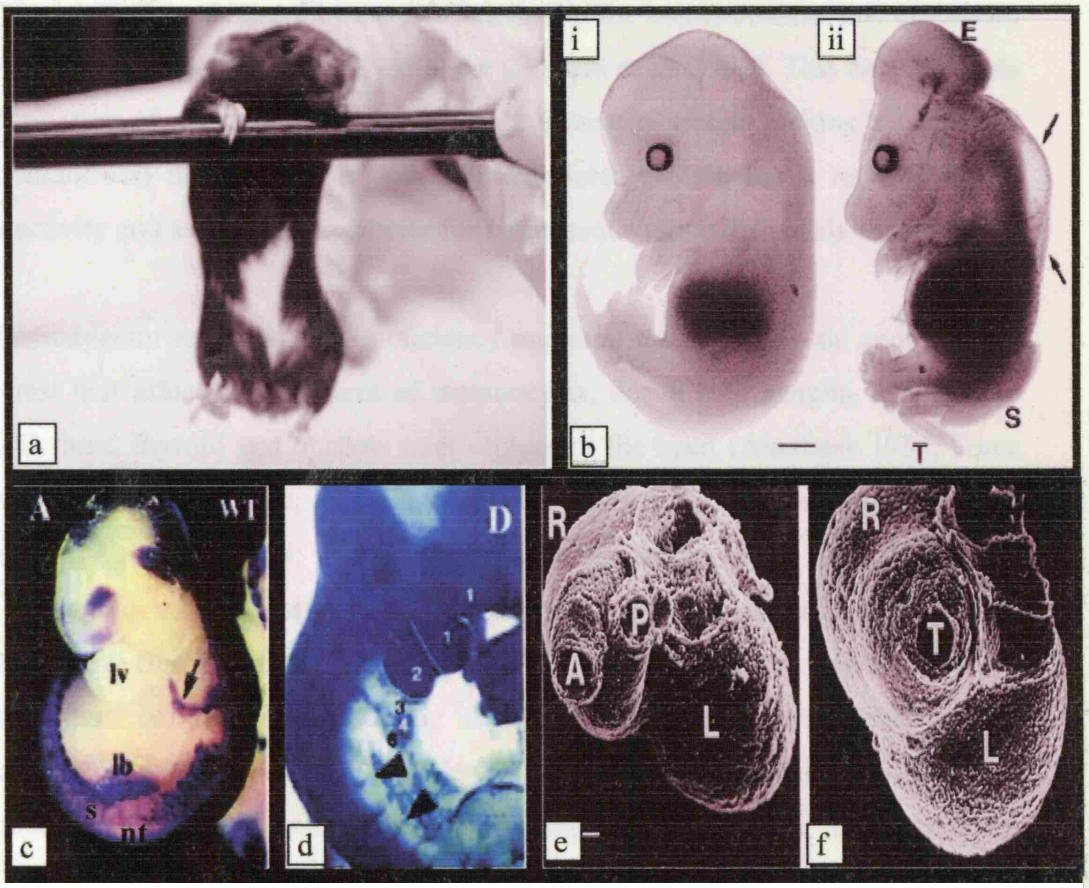


Figure 6.1a: *Pax3* heterozygote mouse with a white belly "splotch"; **bi**: wild type E13.5 mouse embryo; **bii**: *Pax3* homozygous E13.5 embryo with spina bifida (S) and exencephaly (E). Arrows indicate subcutaneous oedema from heart failure. (Conway et al, 1997). **c**: whole mount *in situ* hybridisation of E10.5 wild type embryo showing *Pax3* expression in the dorsal neural tube (nt), somites (s), limb bud (lb) and migrating myoblasts (arrow) (Epstein et al, 2000). **d**: detection of NCCs in the cardiac OFT and pharyngeal arches using a *Cx43-lacZ* transgene. *Sp/Sp* E10.5 embryo was stained with β-galactosidase to detect the *Cx43-lacZ* transgene. Note the small caudal PAA's 3-6. Arrowheads mark NCCs migrating from the dorsal neural tube (Epstein et al, 2000). **ef**: scanning electron microscope from murine embryo hearts at E13.5. L: left ventricle; R: right ventricle, P: pulmonary outflow tract, A: aorta; **e**: a heart from a wild type embryo with normal morphology; **f:** a heart from an embryo with PTA (Conway et al, 1997). **g**: Patient with WSI. Note the white forelock, iris heterochromia and high nasal root (taken from Oxford Medical Database (OMD) with author's permission). **H:** Patient with WSIII. Note the depigmented hair and skin, upper limb muscle atrophy, severe contractures, relative macrocephaly, narrow palpebral fissures, large ears, high nasal root (picture from OMD).

The Sp^{2H} mutation involves a 32 base pair deletion in the homeodomain, one of the two main DNA-binding regions of the *Pax3* gene (Epstein et al. 1991). This deletion leads to the formation of a stop codon resulting in a truncated protein lacking a functional homeodomain and the carboxy region. The translated protein has a reduced DNA-binding activity and an altered specificity for target sequences (Chalepakis et al. 1994a).

In the homozygous state, the *Pax3* deficiency mutation results in several anomalies of neural crest that affect development of melanocytes, dorsal root ganglia, sympathetic ganglia, thymus, thyroid and outflow tract septum of the heart (Auerbach 1954; Franz 1989; Conway et al. 1997c; Epstein et al. 2000b). *Splotch* mice also lack limb musculature (Auerbach 1954; Franz 1993; Goulding et al. 1994), and suffer from defects of neural tube closure such as spina bifida and exencephaly (figure 6.1b) (Auerbach 1954; Beechey 1986; Franz 1989). In addition tissues of somitic origin are affected in *Pax3* deficient mice resulting in defects of the ribs, neural arches and the acromion processes of the scapulae (Henderson et al. 1999). Figure 6.1c demonstrates a whole mount *in situ* hybridisation pattern showing *Pax3* expression at E10.5 in the mouse.

6.1.2.1 Cardiac defects in the Sp^{2H} mouse

The hearts of heterozygous Sp^{2H} embryos are phenotypically normal and these embryos are viable and fertile, whilst 60% of Sp^{2H} homozygote embryos die at midgestation from cardiac defects (Conway et al. 1997c). The remaining Sp^{2H}/Sp^{2H} embryos have normal hearts and die at term from respiratory failure secondary to absence of a muscular diaphragm (Li et al. 1999) and neural tube defects.

As is seen in *Tbx1* homozygous mice and some patients with 22q11DS, 50% of Sp^{2H} homozygous have defective septation of the outflow tract resulting in persistent truncus arteriosis (PTA) (figure 6.1e and f) (Franz 1989; Conway et al. 1997c) and a further 7% have double outlet right ventricle (DORV) (Conway et al. 1997c). An array of abnormalities of aortic arch derived vessels can occur, including aberrant right subclavian artery due to abnormal regression of the right 4th pharyngeal arch artery (Franz 1989). *Pax3* expressing neural crest cells are fated to populate pharyngeal arches 3, 4 and 6 and *Splotch* homozygous embryos have hypoplasia of these caudal arches, similar to the findings seen in *Tbx1* nulls (figure 6.1d) (Epstein et al. 2000b). In addition

abnormal excitation-contraction coupling has been demonstrated in the *Splotch* myocardium (Conway et al. 1997a).

6.1.2.2 Splotch as a mammalian model to study neural crest migration

The Sp^{2H} mutant mouse has served as an important model to study the relationship between neural crest migration and cardiac development. A neural crest cell (NCC) population termed the cardiac neural crest, emerges from the occipital neural tube and migrates through the 3rd, 4th and 6th branchial arches making a cellular contribution to pharyngeal derivatives, including the aortic arch arteries, thymus, thyroid and parathyroid glands.

Studies in chick embryos indicate that disruption of NCC migration can result in cardiac outflow tract (OFT) defects such as failure of septation of the OFT into the aorta and pulmonary artery resulting in PTA, and DORV due to failure of juxtaposition of the aorta with the left ventricle (Kirby et al. 1983). This results from insufficient looping and remodelling of the inner curvature of the heart.

Some experimental groups have supported the theory that *Splotch* represents a mammalian animal model of failed neural crest migration (Moase & Trasler 1990; Conway et al. 1997b). However in 2000 Epstein *et al* utilized a *Cx43-lacZ* transgene marker to label NCCs and showed that cardiac neural crest cells *do* emerge and migrate to the heart OFT in *Sp/Sp* embryos (Epstein et al. 2000b) prompting the conclusion that *Pax3* is not essential for cardiac neural crest migration. Epstein *et al* also demonstrated a reduction in the abundance of β-galactosidase expressing cardiac NCCs in the OFT of *Sp/Sp* mutants and observed that the severity of the deficit correlated with the OFT defect. Thus embryos with PTA at E13.5 had less β-galactosidase expressing NCCs than embryos with DORV. The latter findings were consistent with those of Conway *et al* (Conway et al. 1997b).

6.1.2.3 Delayed onset of neural crest migration in Splotch and NCCs in the OFT

More recently, A delay by in the onset of cardiac neural crest emigration from the Sp^{2H} homozygous neural tube has been detected by 3 somite stages (i.e. approximately 6 hours of development) (Chan et al. 2004). This group also demonstrated that mutant

NCCs are capable of migrating along normal pathways and that there is a significant reduction of cells in Sp^{2H} homozygotes, and a smaller reduction in heterozygotes (Chan et al. 2004).

6.1.2.4 Are the cardiac NCC deficits cell autonomous in Sp^{2H}?

There is much debate in the literature as to whether the NCC defect seen in *Splotch* embryos is autonomous to the NCC lineage or non-autonomous and due to the mutant cellular environment in which the NCCs migrate Rescue of the OFT defects by expression of wild type *Pax3* driven by a neural crest and neural tube specific region on the *Pax3* promotor (Li et al. 1999) provides evidence for a cell autonomous role came from. However Henderson *et al* demonstrated that increased expression of the extracellular matrix protein versican caused defective NCC migration in Sp^{2H} (Henderson et al. 1997) suggesting a non cell-autonomous role for *Pax3*.

To improve understanding of the mechanisms of NCC abnormalities in *Splotch*, Chan *et al* transplanted labelled pre-migratory NCCs between Sp^{2H} embryos of different genotypes and found that the defects of cardiac neural crest migration only occurred when both donor and recipient embryos were of a mutant genotype (Chan et al. 2004). They concluded that defects seen in Sp^{2H}/Sp^{2H} embryos require the genetic defect in both NCCs and their migratory environment (Chan et al. 2004).

6.1.3 PAX3 mutations in human patients: Waardenburg syndrome

Mutations in the human homologue of *Pax3* (paired domain or homeodomain) have also been identified as a genetic cause for Waardenburg syndrome (WS) (Baldwin et al. 1992). In Waardenburg syndrome type 1 (WS1:OMIM 193500) individuals with a heterozygous mutation in *PAX3* display abnormalities such as deafness (in 25%), pigmentary abnormalities (white forelock and white skin patches), characteristic facial features including dystopia canthorum (increased distance between the inner canthi), and iris heterochromia (figure 6.1g). A small number of patients have cardiac abnormalities: usually a VSD. WS2 (OMIM 193510) can be caused by mutations in *Microphthalmia-associated transcription factor (MITF)* (Tassabehji et al. 1994) or *SLUG* (Sanchez-Martin et al. 2002) and is characterised by a higher incidence of deafness (50%) and iris heterochromia, without the facial characteristics seen in WS1.

Table 6.1.3. Summary of human and mouse forms of Waardenburg syndrome detailing the main features in each.

	Human	Mouse	Mouse	Human	Mouse	Mouse	Mouse	Human	Human	Mouse	Mouse	Mouse
	WS1	WS1 +/-	WS1-/-	WS2	WS2 -/-	WS2 -/-	WS2 +/-	WS3	WS4	WS4	WS4	WS4
Cardiac	VSD		OFT									
Deafness	25%	×		50%	✓	?	✓	✓	✓	?		✓
Depigmentation	✓	✓			✓	✓	✓	✓	✓			✓
Eye defect	IH			IH	✓[9]			IH+✓[7]	IH			
NTDs	Rare[2]		✓					Rare				
Cleft lip/palate	3%											
Facial features[1]	✓	✓		×			×	✓	✓			
Gland defects*	×		✓									
Skeletal	Rare[3]		✓					✓[6]				
Hirchsprung disease	Rare			Rare					✓	✓	✓	✓
Muscle	×		✓					✓[5]	✓[5]			
Nerve	×		✓						✓[8]			
Autonomic	×		✓						✓			
Genitalia	Rare[4]											
Gene(s)	*PAX3*	*Pax3*	*Pax3*	*MITF*	*Mi*	*Slugh*	*Ednrb*	*PAX3*	*EDN3*	*Edn3*	*Ednrb*	*Sox10*
				SLUG					*SOX10*			
									EDNRB			

Key

* Thymus/thyroid/parathyroid

1 Dystopia canthorum, synophrys, high nasal bridge, hypoplasia of the alae nasae

2 Meningomyelocele

3 Osseus syndactyly

4 Malformed uterus/vagina/anal atresia

5 Muscle atrophy/contractures

6 Carpal synostosis/cartilaginous/bony exostoses

7 Strabismus/gaze palsy and blepharospasm

8 Case reports of myelin deficiency

9 Microphthalmia and absent pigment epithelia

WS Waardenburg Syndrome

NTD Neural tube defects

OFT Outflow tract

IH Iris Heterocromia

Type 3 WS (WS3: OMIM 148820) is caused by heterozygous and homozygous mutations in *PAX3* (paired domain or homeodomain) and patients have all of the features seen in WS1, but with hypoplasia of the limb muscles (correlating with the complete absence of limb muscles seen in Sp^{2H} homozygote embryos) (figure 6.1h). WS4 is an autosomal recessive (when caused by mutations in *EDNRB* or *EDN3*) or autosomal dominant (when cause by mutations in *Sox10*) disorder characterised by all of the features of WS1 but with colonic agangliosis resulting in megacolon in infancy due to failure of NCC migration into the colon (OMIM 277580). There is neurological variant of WS4 whereby patients can be affected with peripheral neuropathy (consistent with Charcot-Marie-Tooth disease type 1) and a deficiency of brain myelin (Inoue et al. 1999). Some patients can also have autonomic dysfunction suggestive of abnormal sympathetic development which is also abnormal in *Splotch* embryos (Touraine et al. 2000). Mutations in *endothelin-3* (*EDN3*), (Edery et al. 1996), endothelin B receptor gene (*EDNRB*) (Puffenberger et al. 1994) and *Sry-Box 10* (*SOX10*) usually causing the neurological variant (Touraine et al. 2000) have all been reported in WS4. Table 6.1.3 summarises the human and mouse forms of Waardenburg syndrome detailing the main features in each.

6.1.4 Cascade of Waardenburg gene products

Not surprisingly most of the genes involved in the different types of WS are interrelated. Figure 6.1.4 (Tachibana et al. 2003) shows the relationship of the gene products of the six WS genes in melanogenesis: *EDN3, EDNRB, SOX10, PAX3, microphthalmia transcription factor (MITF)* and *SLUG*. Endothelin 3 (EDN3) is the ligand for the endothelin-B receptor (EDNRB). Binding of EDN3 to its receptor is thought to cause elevation of cyclic AMP (cAMP) which results in activation of extracellular signal-related kinase (ERK) which phosphorylates MITF and promotes its degradation (Wu et al. 2000). Activation of glycogen synthase kinase (GSK3β) by cAMP leads to phosphorylation and activation of MITF (Bertolotto et al. 1998). Elevated cAMP also activates protein kinase A (PKA) and phosphorylates cAMP response element binding protein (CREB) which binds the *MITF* promoter. PAX3 (Watanabe et al. 1998) and SOX10 (Bondurand et al. 2000) also bind with the *MITF* promoter and MITF binds and causes activation of the *SLUG* promoter (Sanchez-Martin et al. 2002).

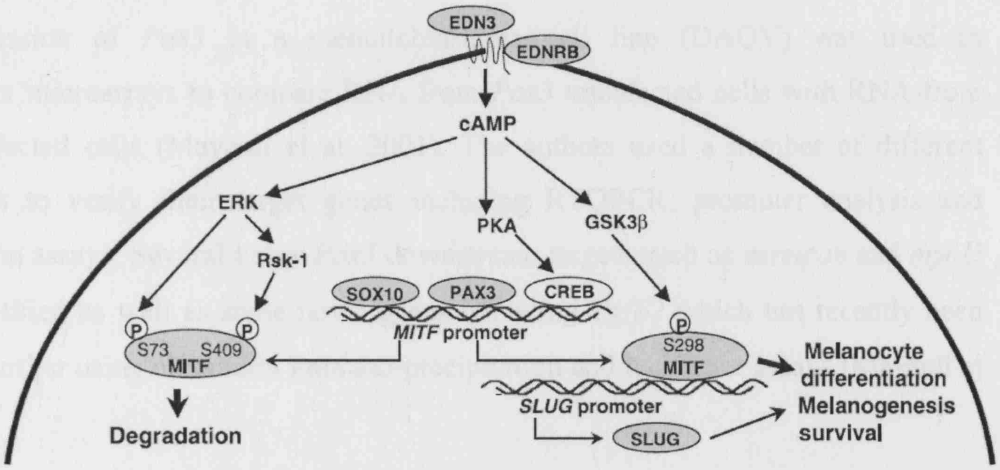


Figure 6.1.4.

Hierarchic relationship of gene products of six Waardenburg syndrome *genes: EDN3, EDNRB, SOX10, PAX3, MITF* and *SLUG* (Tachibana et al. 2003).

6.1.5 Current knowledge of Pax3 expression targets

To date most of the literature on identification of *Pax3* targets has concentrated mainly on muscle related genes such as *MyoD* (Maroto et al. 1997). This is the pathway most studied in *Pax3* presumably because it has implications for the treatment of rhabdomyosarcomas (RMSs), a *PAX3* associated malignancy (see paragraph 6.4.5). A cDNA microarray experiment using *PAX3* and *PAX3-FKHR* (a *Pax3* fusion gene involved with RMSs transfected into NIH3T4 cell lines) identified a number of myogenic targets including upregulation of *myogenin* and *Slug* by *PAX3-FKHR* (Khan et al. 1999). *PAX3* was found to act as a repressor and decay accelerating factor for complement (*Daf), paired mesoderm homeobox 1 (Pmx1)* and *Platelet derived growth factor α (Pdgfrα)* were suggested as potential repressed targets (Khan et al. 1999).

The latter point is noteworthy as several authors have suggested that Pax3 can act as a transcriptional repressor as well as an activator (Chalepakis et al. 1994b). One such example of this is the ability of Pax3 to cause repression of *Msx2* expression via a direct effect on the Pax3 binding site in the *Msx2* promotor (Kwang et al. 2002). Consistent with this, loss of *Msx2* function rescued the cardiac defects of *Splotch* mutant embryos (Kwang et al. 2002).

Overexpression of *Pax3* in a medulloblastoma cell line (DAOY) was used in Affymetrix microarrays to compare RNA from *Pax3* transfected cells with RNA from non-transfected cells (Mayanil et al. 2001). The authors used a number of different techniques to verify their target genes including RTQPCR, promoter analysis and transfection assays. Several know *Pax3* downstream targets such as *versican* and *myoD* were identified as well as some novel genes including *Tgfβ2* which has recently been verified further using chromatin immuno-precipitation and luciferase assays (Mayanil et al. 2006).

6.2 Methodology for identification of *Pax3* target genes

6.2.1 Affymetrix expression microarray study

Please refer to the chapter 2 (paragraph 2.3.4) for the full methodology used for the Affymetrix microarray study. Chapter 4 presents a thorough description of Affymetrix microarray technology and the Genechip system. Paragraph 4.4.1 describes the different meanings of the word "target".

As the current literature appeared to concentrate on muscle related targets of *Pax3*, this study examines the targets relating to neural crest migration and pharyngeal development. For this reason the same branchial arch dissections (see figure 2.3.4.3) were performed for E10.5 wild type, homozygote and heterozygote Sp^{2H} mouse embryos as were used to examine *Df1* downstream targets. Yolk sacs were taken for genotyping. Total RNA was extracted from the embryos using TRIzol. RNA was tested on the Agilent bioanalyser and total RNA with poor yield or degradation was rejected. 3-4 embryo total RNAs were pooled for hybridisation to each Affymetric MG-U74Av microarray chip. cDNA was generated from the total RNA extracted and this was transcribed and biotinylated to form cRNA. Three chip hybridisations were done for the wild type and homozygous Sp^{2H}/Sp^{2H} embryos and two for the heterozygous $Sp^{2H}/+$ pools to reduce costs.

6.2.1.1 Genespring statistical filtering

Metric text files were taken from the Affymetrix analysis files and imported into Genespring. The signal intensities of all wild type chips were normalised to 1 to allow comparison with genes changed in homozygote and heterozygote chips. Filtering in Genespring was done using the data interpretation of "mutant samples treated as separate" model shown in figure 4.3.5.1. If the "mutant samples averaged" model was used then too many genes passed the filtering methods implying the first method is more stringent.

The first constraint applied was to filter out genes if they had very low expression levels or control strengths. The next constraint was expression level. As this experiment included homozygote embryos the 0.5 cut-off value of downregulation was used as a more stringent boundary than the 0.75 which was used for the heterozygous *Dfl* embryos.

Statistical filtering was always modified to allow for the most stringent filtering down to reasonable numbers. For the upregulated gene test the parametric Students' T test (assume variances equal) was used. For the downregulated list all of the genes passed most types of statistical analysis (see figure 6.3.2 in the results section of this chapter). A detection p cut off value was not used for this dataset as too few genes passed through the statistical filtering with this method.

6.2.2 Validation of potential targets

Please refer to the chapter 2 and chapter 5.2 for the full methodology used for validation of potential targets by using RTQPCR and *in situ* hybridisation.

6.2.2.1 Real Time Quantitative PCR

Chapter 5 presents a thorough description of the technology employed by the RTQPCR techniques used in this study.

To summarise, total RNA was extracted from the branchial arch region of E10.5 mouse embryos and embryos were genotyped by their yolk sac DNA. Three-four embryos at 30-34 somite stages were combined for wild type, homozygous and heterozygous Sp^{2H}

embryos (validation was done with different total RNA to that used for Genechip hybridisation). RNA was discarded if either the yield or quality were below expected levels. Total RNA was reverse transcribed to cDNA using random hexamer primers. The resulting cDNA was diluted 1:50 to use in a real time quantitative PCR reaction (RTQPCR) reaction using SYBRGreen (QIAgen) for the detection of fluorescence during amplification. GAPDH was used as a control as there was no difference in GAPDH expression levels between wild type and Sp^{2H} mutant cDNA on the microarrays. All reactions were performed in triplicate in a DNA Engine Opticon 2 system (M. J. Research inc.) and a threshold cycle (Ct) value of 0.001 was used for analyses. RTQPCR experiments were performed with cDNA at least twice and relative ratios were averaged. The relative expression ratios were arrived at by using the $2^{-\Delta\Delta Ct}$ method of data analyses (Livak & Schmittgen 2001) with GAPDH levels used to standardise the cDNA input quantities.

6.2.2.2 Whole mount In situ hybridisation

Constraints on the number of *Pax3* embryos available meant it was only possible to perform a whole mount *in situ* hybridisation experiment with one probe. The probe used was *Wnt5a* and was obtained from Dr D Henderson from the Neural Development Unit, Institute of Child Health. The experiments were performed more than once with 3-4 embryos used in each experiment and a sense probe used as a control.

6.3 Results

6.3.1 Embryo genotyping and RNA analysis

In common with the *Df1* experiment (chapter 4) only embryo yolk sacs that gave an unequivocal PCR result on genotyping were included in the analysis. RNA samples with a poor quality and yield were excluded from analysis (see figure 4.3.2 for examples).

6.3.2 Affymetrix Microarray Hybridisation

MG-U74Av Affymetrix microarray chips were used for hybridisation and chips with poor hybridisation parameters were rejected. Table 6.3.1 shows a summary of the report

files which show the hybridisation parameters for each chip. Interpretation of these parameters has already been discussed previously for the *Df1* experiment in chapter 4 (paragraph 4.3.3). Three chips were rejected due to a high background (Pax3Wt2, Pax3Het4, Pax3Homo4) and two chips were rejected due to a high scale factor (Pax3Het1 and Pax3Het3).

Chips with acceptable parameters and therefore included in data analysis:

	RNA Type	SF	Raw Q	BG	%P	%A	3'/5'(GAPDH)
Pax3Wt1	3 Pax3 Wt embryos	3.4	2.4	65	33.9	63.8	1.46
Pax3Wt3	3 Pax3 Wt embryos	1.7	2.46	70	45.2	52.6	1.1
Pax3Wt4	3 Pax3 Wt embryos	10.4	1.7	39	28	60.7	1.89
Pax3Homo1	3 Pax3 Homo embryos	2.45	2.55	71	36.7	60.7	2.93
Pax3Homo2	4 Pax3 Homo embryos	3.18	2.04	52	38.3	59.2	1.67
Pax3Homo3	5 Pax3 Homo embryos	7.66	1.75	29.8	29.8	68	1.22
Pax3Het2	3 Pax3 Het embryos	1.12	1.97	45	51.6	46.4	1.04
Pax3Het5	3 Pax3 Het embryos	2.45	1.95	44.9	45.2	52.4	1

Chips rejected from data analysis due to unacceptable quality parameters:

	RNA Type	SF	Raw Q	BG	%P	%A	3'/5'(GAPDH)
Pax3Wt2	3 Pax3 Wt embryos	1.6	7.75	385	15.4	83	1.2
Pax3Het1	3 Pax3 Het embryos	24.8	1.4	33	20.6	77.2	1.1
Pax3Het3	4 Pax3 Het embryos	31.4	1.5	32	14	84	1.6
Pax3Het4	5 Pax3 Het embryos	1.4	7.8	373	15.6	82.6	1.1
Pax3Homo4	4 Pax3 Homo embryos	1.27	15.6	1125	2.2	97.4	0.9

Table 6.3.1

Summary of the report files of the *Pax3* microarray chips used in the data analysis. **SF**: scale factor, **BG:** background intensity, **%P**: % genes called as present, **%A**: % genes called as absent, **3'/5' GAPDH**: ratio of 3' and 5' transcripts of GAPDH used as a measure of degradation. Shaded samples were rejected for analysis.

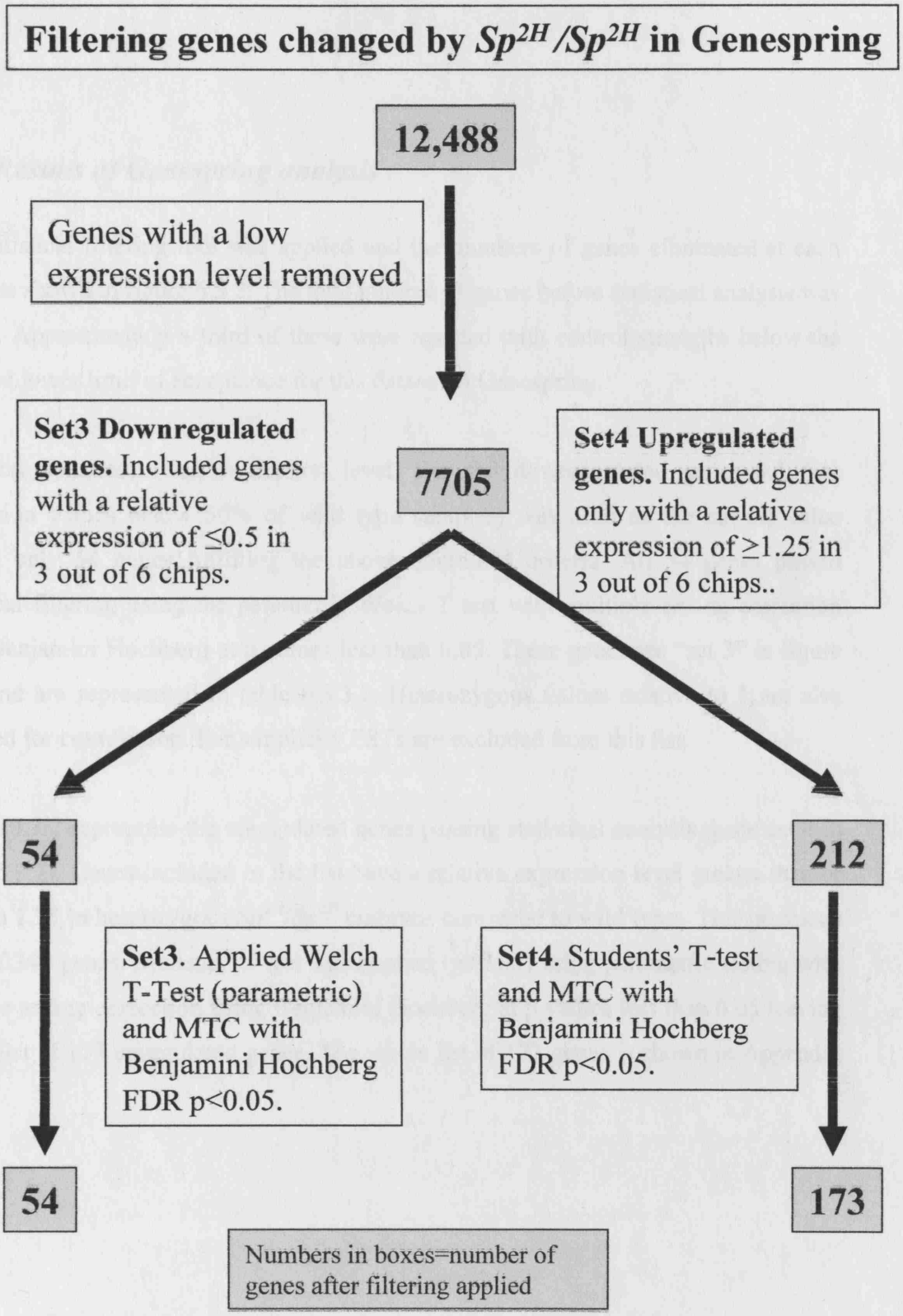


Figure 6.3.2
Diagram representing the number of genes eliminated at each stage of filtering in Genespring in the Sp^{2H}/Sp^{2H} dataset where wild type chip data is normalised to 1.
MTC: multiple testing correction; FDR: false discovery rate.

6.3.3 Results of Genespring analysis

The statistical filtering tool was applied and the numbers of genes eliminated at each stage are shown in figure 6.3.2. The total number of genes before statistical analysis was 12,488. Approximately a third of these were rejected with control strengths below the assigned lower limit of acceptance for this dataset in Genespring.

The next constraint was expression level. For the downregulated genes 0.5 (i.e. expression values below 50% of wild type samples) was used as the cut-off value leaving only 54 genes fulfilling the above combined criteria. All 54 genes passed statistical filtering using the parametric Welch T test with multiple testing correction using Benjamini Hochberg at p values less than 0.05. These genes are "set 3" in figure 6.3.2 and are represented in table 6.3.3.1. Heterozygous values relative to 1 are also provided for comparison. For simplicity ESTs are excluded from this list.

Table 6.3.3.2 represents the upregulated genes passing statistical analysis (gene set 4 in figure 6.3.2). Genes included in the list have a relative expression level greater than or equal to 1.25 in homozygous Sp^{2H}/Sp^{2H} embryos compared to wild type. This produced a list of 340 genes. Students' T test was applied ($p<0.05$) using parametric testing with multiple testing correction using Benjamini Hochberg at p values less than 0.05 leaving a final list of 173 upregulated genes. The whole list of 173 genes is shown in Appendix 2.

Table 6.3.3.1 *Pax3* Downregulated genes. Data set 3.

Affymetrix No	Transcription factors	Homo	Het
99390	Wnt5a*	0.2	2.1
93708	Protein inhibitor of activated STAT3 (PIAS3)*	0.3	1.4
94102	Homeobox Hmx1*	0.4	1.6
99464	methyl CpG binding protein 1	0.4	1.8
103009	Hira*	0.4	1.5

Affymetrix No	Cell cycle regulators	Homo	Het
101465	Stat1	0.3	0.7
101969	neuroblastoma, suppression of tumorigenicity 1	0.3	1.1
92626	neural proliferation, differentiation and control gene 1	0.3	1.3
98499	caspase 7	0.3	1.4
102753	multiple endocrine neoplasia 1	0.3	1.8

Affymetrix No	Extracellular Matrix Proteins	Homo	Het
160583	extra cellular link domain-containing 1	0.2	0.7
104205	aggrecan 1*	0.4	0.2

Affymetrix No	Receptors/Transmembrane proteins	Homo	Het
161689	interleukin 1 receptor, type II	0.3	0.3
103973	potassium inwardly-rectifying channel, subfamily J, member 1	0.3	0.4
92699	solute carrier family 7 member 9 (amino acid transporter)	0.3	0.6
98332	potassium inwardly-rectifying channel, subfamily J, member 9	0.3	0.9
100759	mannose receptor c type 2	0.3	1
93151	copine VI	0.4	0.6

Affymetrix No	Cytoskeletal proteins	Homo	Het
94379	kinesin family member 1B	0.2	0.9
92462	Septin 6	0.3	0.4
93178	neuronal guanine nucleotide exchange factor (ephexin)	0.3	0.6

Affymetrix No	Enzymes	Homo	Het
102285	O-6-methylguanine-DNA methyltransferase (DNA repair)	0.2	0.6
94843	polymerase (DNA-directed), delta 4	0.2	0.7
93851	Rab geranylgeranyl transferase, a subunit	0.2	1.6
161021	p21 activated kinase 3*	0.3	0.6
95015	aldo-keto reductase family 1, member C12	0.3	0.6
103072	hydroxysteroid dehydrogenase-1, delta<5>-3-beta	0.3	0.7
160601	lunatic fringe*	0.3	0.8
94197	UDP-glucose ceramide glucosyltransferase	0.3	0.9
102840	protein tyrosine phosphatase, non-receptor type 12	0.4	1.3

Affymetrix No	Uncategorised proteins	Homo	Het
103639	interferon-induced protein with tetratricopeptide repeats 2	0.3	0.6
160100	EF hand domain containing 2	0.3	0.7
93630	CUG triplet repeat, RNA binding protein 1	0.3	1.6
103286	signal transducing adaptor molecule (SH3 domain and ITAM motif) 2	0.3	1.6
95311	HLA-B associated transcript 2	0.4	0.3
98293	G substrate	0.5	1.2

*followed up with RT-PCR

Homo: relative expression on microarray in Sp^{2H}/Sp^{2H} (homozygous) embryos

Het: relative expression on microarray in $Sp^{2H}/+$ (heterozygous) embryos

Cited in discussion as potentially biologically relevant

Figures show relative expression of genes with wild type expression levels normalised to 1

Table 6.3.3.2 *Pax3* Upregulated genes. Data set 4.

Affymetrix No	Cell adhesion/migration	Homo	Het
97375	polycystin-1	2.2	2.6
98798	chondroitin sulfate proteoglycan 3 (neurocan)	2.2	1.7
92702	astrotactin 1	1.9	2
99492	catenin beta interacting protein 1	1.8	1.6
99007	flotillin 2:cell adhesion	1.7	2.5
103729	laminin, alpha 1	1.6	2.1
101650	protocadherin alpha 6	1.6	1.1
161786	integrin beta 5	1.5	1.8
102261	procollagen, type XIII, alpha 1	1.5	1.4
161535	polycystin-1	1.3	1.4

Affymetrix No	Cell cycle regulators	Homo	Het
	apoptosis promotors		
93231	hic-1	1.9	1.3
102919	fibrosarcoma oncogene family, protein K (avian)	1.8	1.3
102963	E2F transcription factor 1	1.6	2
98031	Bcl-2-related ovarian killer protein (Bok)	1.6	1.1
161500	Bcl-associated death promoter (BAD)	1.5	1.1
103514	tumor necrosis factor receptor superfamily, member 21	1.5	1.9
	anti-apoptosis		
160895	Fanconi anaemia complementation group C	2	1.5
104170	mitogen activated protein kinase 8 interacting protein	1.9	1.9
160920	Bcl2-like 2	1.7	1.8
99100	Stat3	1.5	1

Affymetrix No	Transcription factors	Homo	Het
103735	Wnt6*	2.3	2.7
101395	purine rich element binding protein B	1.9	1.7
161030	Scleraxis (cartilage and bone formation)	1.9	1.6
102411	zinc finger protein 239	1.8	2.6
104568	myeloid/lymphoid or mixed-lineage leukemia	1.7	1.9
94709	Gcm2	1.7	1
98808	neurogenic differentiation 2	1.6	0.7
101684	M.musculus T2 mRNA.	1.5	2.7
103651	general transcription factor IIF, polypeptide 2	1.4	1.6
92914	HoxB7	1.4	0.9
92979	E1A enhancer binding protein, E1AF/ PEA3	1.3	1.6
100467	lymphoblastomic leukemia	1.3	1.3
97660	Mus musculus mRNA for TCF-4 protein.	1.3	1.3

Affymetrix No	Cytoskeletal proteins	Homo	Het
161545	myosin, heavy polypeptide 4, skeletal muscle	2.1	1.8
95466	coactosin-like 1 (Dictyostelium)	1.7	3.3
104255	ESTs, Weakly similar to Diaphanous-related formin 3	1.6	2.5
92635	tubulin, alpha 4: cytoskeletal	1.5	1.4
98126	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	1.5	0.5
161990	diaphanous homolog 1 (Drosophila)	1.4	1.7
100440	Ankrin1	1.4	1.6
102426	calsequestrin 1	1.4	1.3
93266	tropomyosin 3, gamma	1.3	1.1

*didn't pass original statistical filter

Homo: relative expression on microarray in Sp^{2H}/Sp^{2H} (homozygous) embryo

Het: relative expression on microarray in $Sp^{2H}/+$ (heterozygous) embryos

Cited in discussion as potentially biologically relevant

Figures show relative expression of genes with wild type expression levels normalised to 1.

Table 6.3.3.2 *Pax3* Upregulated genes. Data set 4.

Affymetrix No	Signal transduction	Homo	Het
162204	Notch gene homolog-1†	2	2.3
161348	PDZ and LIM domain 1 (elfin)	2	1.1
98486	stathmin-like 4	1.9	1.8
160409	phosphatidylinositol transfer protein	1.8	1.9
93456	Bone morphogenetic protein 4†	1.7	1.8
97151	orphan G protein-coupled receptor; Mus musculus	1.7	0.9
160629	regulator of G-protein signalling 10	1.6	1.7
92302	Son of sevenless homolog 2 (Drosophila)	1.5	2.2
98278	sprouty homolog 4 (Drosophila)	1.4	1.1
96099	casein kinase II, beta subunit	1.3	0.9

Affymetrix No	Receptors/Transmembrane proteins	Homo	Het
101131	cholinergic receptor, nicotinic, alpha polypeptide 7	3.4	0.9
97053	single cysteine motif 1 receptor	2.1	1.2
104680	receptor (calcitonin) activity modifying protein 1	2	1.9
93018	interferon induced transmembrane protein 3-like	1.9	1.2
92235	retinoid X receptor alpha	1.8	1.6
99399	Mus musculus agouti-related protein (Agrp) gene	1.8	1
100327	precursor; Mouse LT lymphotoxin (LT) gene	1.8	1
160134	adiponectin receptor 1	1.6	1.8
100282	5-hydroxytryptamine (serotonin) receptor 7	1.6	1.3
92437	transmembrane 7 superfamily member	1.5	1.6
95104	syndecan 2	1.5	1.6
96174	nuclear pore membrane protein 121	1.4	2.7
97224	proline-rich nuclear receptor coactivator 1	1.4	1
97966	nuclear receptor coactivator 5	1.3	1.6

Affymetrix No	Uncategorised Proteins	Homo	Het
99847	sialyltransferase 4A	2.1	1.9
160869	sirtuin 3 (S. cerevisiae)	1.4	2.6
103552	Rabaptin	1.7	2.7
96310	myelin basic protein	1.7	1.4
100617	Slc25a5 (Ant2)	1.7	0.9
97821	p21 activated kinase 2†	1.6	1.8
98926	vesicle associated membrane protein 2	1.4	2.2
97983	Syntaxin binding protein 1	1.4	2.1

†didn't pass original statistical filter

Homo: relative expression on microarray in Sp^{2H}/Sp^{2H} (homozygous) embryo

Het: relative expression on microarray in $Sp^{2H}/+$ (heterozygous) embryos

Cited in discussion as potentially biologically relevant

See Appendix 2 for full list of upregulated genes.

Figures show relative expression of genes with wild type expression levels normalised to 1.

6.3.4 Validation results

6.3.4.1 Real Time quantitative PCR (RTQPCR)

Due to time restraints it was only possible to follow-up a few potential differentially expressed genes in this experiment.

Affymetrix No.	Gene	MA Homo	RTQPCR Homo	MA Het	RTQPCR Het
Not on array	Pax3	NA	0.3	NA	0.7
93708	Protein Inhibitor of activated Stat3 (PIAS)	0.3	0.5	1.4	0.6
104205	Aggrecan 1	0.4	0.5	0.2	0.5
99390	Wnt5a	0.2	0.5	2.1	0.5
94102	Hmx1	0.4	0.4	1.6	0.6
160601	Lunatic Fringe	0.3	0.6	0.8	1.0
103009	Hira	0.4	0.8	1.5	0.8
161021	p21 activated kinase 3 (Pak3)	0.3	1.1	0.6	1.0
94709	Gcm2	1.7	0.8	1.0	1.1

Table 6.3.4.2

Results of verification experiments using RTQPCR to analyse the *Pax3* microarray results. Numbers shown are expression levels relative to a wild type value of 1. KEY: MA: microarray, Homo: homozygous Sp^{2H}/Sp^{2H}; Het: heterozygous $Sp^{2H}/+$ embryo.

6.3.4.2 Whole mount In situ hybridisation

Figure 6.3.4.2 shows a representation of wild type (a,b), $Sp^{2H}/+$ (c,d) and Sp^{2H}/Sp^{2H} (e,f) embryos hybridised with the *Wnt5a* riboprobe. The embryos hybridised with the sense probe had background signal (g) only. There is artefactual dye trapping in the heads of the wild type embryos. The probe for *Wnt5a* showed expression in the distal aspect of the forelimb and the hindlimb (former>latter); the outflow tract of the heart; the anterior aspect of the somites and a dense area between the forelimb bud and the branchial arches, posterior to the heart, which is likely to be the lateral plate mesoderm. This area labelled with a white arrow "↑" gave a weaker signal in the $Sp^{2H}/+$ and Sp^{2H}/Sp^{2H} embryos compared to the wild type embryos. Cardiac outflow tract staining was too weak to compare.

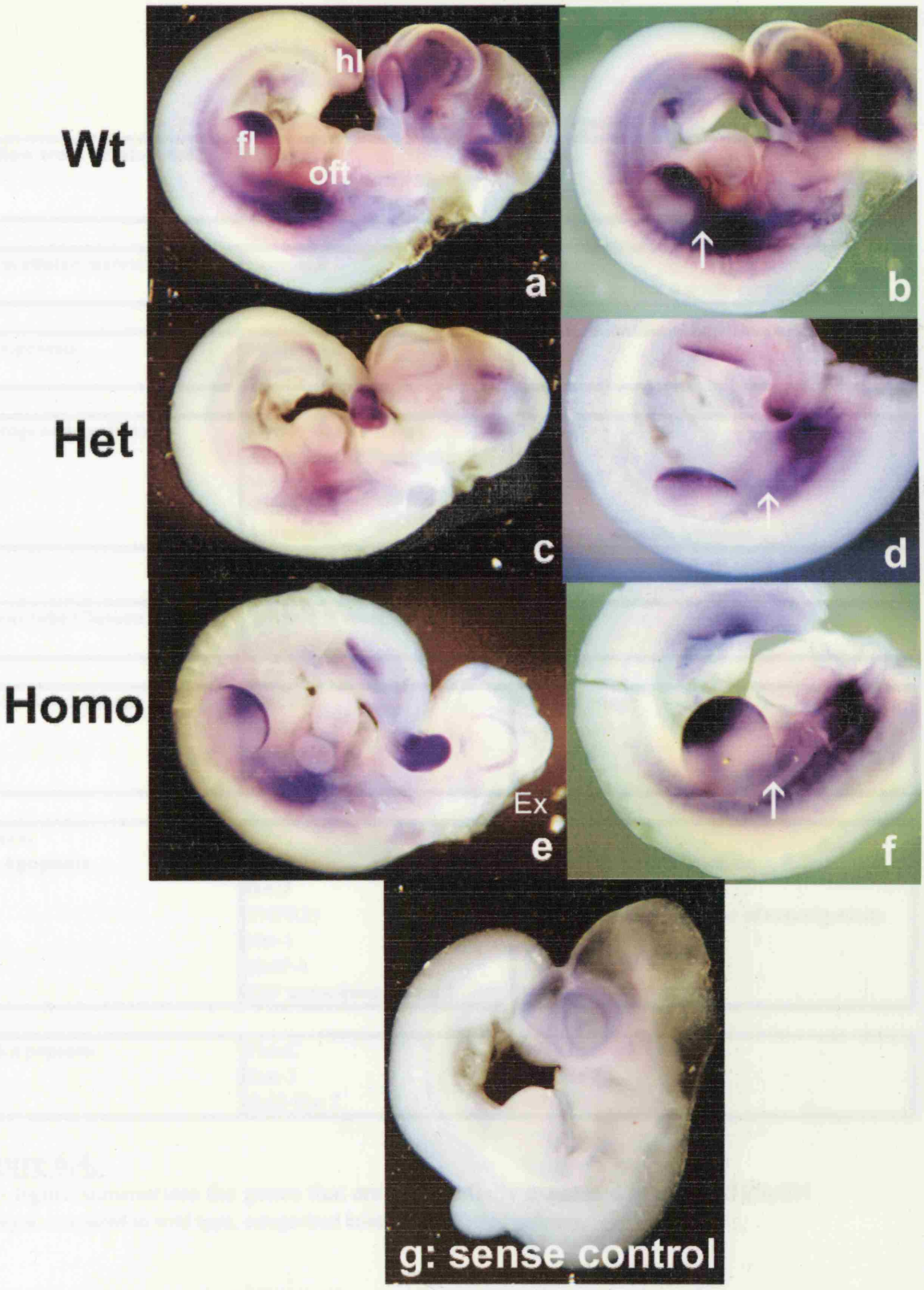


Figure 6.3.4.2

Whole mount *in situ* hybridisations using *Wnt5a* as a probe (E10.5). Left lateral views.

a,b: Wild type antisense (AS); c,d: *Pax3* +/- AS

e,f: *Pax3* -/- AS; g: Wild type sense control

fl: forelimb bud: hl: hindlimb bud;oft: outflow tract;

Ex: Exencephaly;↑: lateral plate mesoderm

	Upregulated	*Downregulated*
Outflow tract development	Polycystin 1 Retinoid X receptor alpha BMP-4	Wnt5a
Extracellular matrix	Neurocan	Aggrecan Extra-cellular link domain-containing 1
Angiogenesis	Syndecan 2 Sprouty 4	
Neurogenesis pathway	Myelin Basic Protein Neurogenic differentiaion 2 MafK Neurocan Astrotactin	Ephexin Kif1B Calnexin NPDC 1 Glucosyl Ceramide Synthase
Neural tube Closure	Retinoid X receptor alpha Hic-1	
Somite development	Notch 1 Pkd1 Scleraxis Wnt 6	Lunatic Fringe
General **Pro-Apoptosis**	Bok BAD TNFR21 Hic-1 BMP-4 E2F transciption factor 1	Glucosyl Ceramide Synthase MEN1 neuroblastoma, suppressor of tumorigenicity PIAS3
Anti-Apoptosis	FancC Stat-3 Bcl2-like 2	NPDC 1 Caspase 7 Stat-1

Figure 6.4.

This figure summerises the genes that are differentially expressed in the Sp2H/Sp2H embryos compared to wild type, categorised broadly by affected pathway.

6.4 Discussion of gene expression changes found in *Splotch* Embryos (see figure 6.4)

Verification using RTQPCR supported the microarray findings for 6/8 of the genes analysed using RNA from homozygous embryos (table 6.3.4). The two genes where this wasn't the case were the potentially downregulated gene *Pak3* (with a RTQPCR of 1.1 in homozygous embryos) and the potentially upregulated gene *Gcm2* (RTQPCR ratio 0.8 in Sp^{2H}/Sp^{2H}embryos). Microarray data was highly variable between homozygous and heterozygous hybridisations. This is because the statistical filtering was performed on the homozygous data and gene expression data is often much more variable in the heterozygous system. There are also many inconsistencies between microarray and RTQPCR data within the heterozygous dataset, with only *Gcm2* and *aggrecan* showing consistent data.

6.4.1 Comparison of microarray results with published data

Comparison of Pax3 microarray results with targets from the two published microarray papers did not reveal any common targets (Khan et al. 1999), (Mayanil et al. 2001). This is not surprising given the methodology used was very different, with neither paper using *Pax3* deficient embryos.

Table 6.4.1 Microarray results for published targets of *Pax3*.

* passed statistical filtering to appear in data set 3. MA Homo: relative expression on microarray in Sp^{2H}/Sp^{2H} embryos. MA het: relative expression on microarray in $Sp^{2H}/+$ embryos.

Affymetrix no	*Pax3* published target	Reference	MA Homo	MA Het
101828	Ret proto oncogene	(Lang & Epstein 2003)	0.38	0.45
160483	MITF	(Watanabe et al. 1998)	0.76	1.61
97955	Tyrosinase related protein 1	(Galibert et al. 1999)	0.35	0.34
95079	PDGF α	(Corry & Underhill 2005)	0.85	1.63
160601	Lunatic Fringe*	(Schubert et al. 2001)	0.3	0.8

Inspection of microarray data for published *Pax3* targets did show "downregulation" in mainly homozygous RNA, although only *lunatic fringe* passed the statistical filtering applied (table 6.4.1).

6.4.2 Genes implicated in Outflow Tract (OFT) development and angiogenesis

6.4.2.1 Wnt5a is downregulated in Sp^{2H}/Sp^{2H}

Pax3 null mice have defects of the OFT such as persistent truncus arteriosus (PTA) and double outlet right ventricle (DORV). Microarray data, RTQPCR data (6.3.4.1) and *in situ* hybridisation (figure 6.3.4.2) experiments have shown that *Wnt5a* is downregulated in *Pax3* homozygous and heterozygous E10.5 embryos (the latter only shown with RTQPCR data).

Wnt5a is a member of the Wnt family of secreted glycoproteins that play essential roles in development. It is a secreted ligand that interacts with the transmembrane receptor *Frizzled2* at the cell surface (Toyofuku et al. 2000). *Wnt5a* and *Frizzled2* have been detected in the OFT of wild type mice at the time of OFT septation, suggesting a possible role in the formation of the cardiac OFT (Toyofuku et al. 2000). OFT abnormalities have been identified in *Wnt5a* mutant mice such as PTA, coarctation of the aorta, IAA and DORV (Schleiffarth et al. 2005). Chick experiments using beads coated with *Wnt5a* showed that ectopic *Wnt5a* protein disrupts normal neural crest migration into the branchial arches (Schleiffarth et al. 2005). This data suggests *Wnt5a* is needed for proper migration of the cardiac NCCs into the cardiac OFT and may function as a migratory cue for cardiac NCCs (Schleiffarth et al. 2005).

Downregulation of *BMP4* and upregulation of *stromelysin-1* has been observed in a microarray study using a cell line overexpressing *Wnt5a* and *frizzled 2* (Prieve & Moon 2003). *Stromelysin-1*, also known as *MMP3*, is an extracellular matrix-degrading metalloproteinase that is involved in cell migration. A study in chick embryos demonstrated that inhibition of matrix metalloproteinases decreases early cardiac neural crest migration (Cai & Brauer 2002). It is therefore possible that downregulated *Wnt5a* in Sp^{2H}/Sp^{2H} embryos contributes to the cardiac phenotype by downregulating *stromelysin-1* and inhibiting cardiac NCC migration.

6.4.2.2 Polysystin-1 (Pkd1) and RXRα are upregulated in Sp^{2H}/Sp^{2H}

Polyscystin-1 (*Pkd1*) is a gene mutated in autosomal dominant polycystic kidney disease (OMIM +601313). It is significantly upregulated in Sp^{2H}/Sp^{2H} (average 1.8 fold) and $Sp^{2H}/+$ (average 2 fold) as determined by two independent probes in this microarray differential expression analysis (table 6.3.3.2). Time constraints did not allow verification with further techniques.

Pkd1 encodes a membrane protein that interacts with *Pkd2* to produce a calcium permeable non selective cation channel (Hanaoka et al. 2000). This suggests that extracellular matrix signals can be transduced by the polycystin complex to regulate diverse cellular processes (Hanaoka et al. 2000). *Pkd1* is strongly expressed in the developing OFT and mice carrying a targeted mutation of *Pkd1* die *in utero* and have structural heart abnormalities such as DORV, PTA, abnormal atrio-ventricular septation and a disorganised myocardium (Boulter et al. 2001).

As *Pkd1* is known to be expressed in NCCs and their derivatives (Guillaume & Trudel 2000) its apparent "upregulation" in Sp^{2H}/Sp^{2H} embryos that have reduced numbers of cardiac NCCs (Chan et al. 2004) is noteworthy. It is possible that this is a compensatory effect for the lack of *Pax3* and that the two genes function in overlapping developmental pathways. One such pathway could be the canonical Wnt pathway (Peifer & Polakis 2000). Wnt signalling plays a pivotal role in NCC population expansion (Ikeya et al. 1997) and cardiogenesis (Marvin et al. 2001) and *Wnt-1* expression was shown to be reduced in the cardiac NCCs of Sp^{2H} embryos (Conway et al. 2000). *Tcf-4* and *β-catenin interacting protein 1* (*ICAT*) are upregulated in Sp^{2H}/Sp^{2H} embryos in this microarray experiment. This may be of molecular relevance as *in vitro* experiments have determined that murine *ICAT* negatively regulates Wnt signalling via inhibition of the interaction between β-catenin and TCF-4 (Tago et al. 2000). *Pkd1* is a known upregulated target of the canonical Wnt pathway (Rodova et al. 2002) and the C terminal portion of *Pkd1* is known to protect soluble β-catenin from degradation (Kim et al. 1999). Administration of pioglitazone (a peroxisome proliferator activated receptor γ [PPARγ] agonist that increases β-catenin levels) to *Pkd1* -/- embryos ameliorated the cardiac defects seen in these mice (Muto et al. 2002).

Gene targeted *RXRα* -/- embryos display defects of the OFT almost identical to those seen in *Pkd1*-/- and Sp^{2H}/Sp^{2H} mice (Gruber et al. 1996). PPARγ forms heterodimers with *RXRα*, which bind to PPAR-responsive elements in the promoters of PPAR target genes resulting in increased β-catenin levels. Interestingly *RXRα* is also significantly upregulated in Sp^{2H} embryos in this experiment (1.8 fold in nulls and 1.6 fold in heterozygous *Pax3* embryos). Since pioglitazone treatment improved the cardiac phenotype in *Pkd1*-/- embryos, it is possible that *Pkd1* and *RXRα* are upregulated in Sp^{2H}/Sp^{2H} to compensate for the loss of *Pax3*.

6.4.2.3 BMP4 is upregulated in Sp^{2H}/Sp^{2H}

BMP4 expression was increased in this microarray experiment with a ratio of 1.7 in Sp^{2H}/Sp^{2H} embryos and 1.8 in $Sp^{2H}/+$ cDNA. The *BMP4* signalling molecule is expressed in OFT myocardium and ventral splanchnic and branchial arch mesoderm suggesting a role for *BMP4* in OFT development. A conditional null allele of *BMP4* driven by *Nkx2.5cre* produced abnormal development of the pharyngeal arteries including interrupted aortic arch type B and PTA (Liu et al. 2004). It is possible that Sp^{2H}/Sp^{2H} embryos upregulate *BMP4* to compensate for lack of *Pax3*, by similar mechanisms to *Pkd-1* and *RXRα*. A further line of investigation might be to generate transheterozygotes between $Sp^{2H}/+$ and either *Pkd-1*, *RXRα* or *BMP4* deleted mice, to produce compound mutant offspring to establish if the cardiac phenotype in Sp^{2H}/Sp^{2H} embryos is rescued.

6.4.2.4 Hira is downregulated in Sp^{2H}/Sp^{2H}

Hira was found to be downregulated to 40% of wild type levels in Sp^{2H}/Sp^{2H} embryos and this downregulation was confirmed by RTQPCR albeit to a lesser degree (table 6.3.4.2). HIRA is known to encode a nuclear protein with histone binding properties that have been conserved from yeast to humans (De Lucia et al. 2001) and may be involved with cell cycle regulation (Hall et al. 2001). Recent work has suggested its requirement for nucleosome assembly independent of DNA synthesis involving histone H3.3 (Loppin et al. 2005; Ray-Gallet et al. 2002). It is an interesting gene for many reasons. Firstly it has been shown to interact with PAX3 in a yeast two hybrid experiment (Magnaghi et al. 1998). Secondly, it is within the DiGeorge syndrome critical region on chromosome 22q11.2 (Scambler et al. 1988), although it is not deleted

in *Df1* mice. Thirdly *Hira* is expressed in the neural tube, the neural crest and neural crest derived structures of the head and branchial arches (Wilming et al. 1997). And finally *Hira* has been found to be required for cardiac OFT septation. When treated with antisense oligonucleotides designed to *Hira* mRNA, a significant portion of treated chick embryos developed PTA (Farrell et al. 1999).

The interpretation of the decreased expression of *Hira* in Sp^{2H}/Sp^{2H} embryos is likely to be partly a reflection of reduced NCC numbers. However given the *in vitro* interaction with PAX3 (Magnaghi et al. 1998), establishment of *Pax3/Hira* mouse crosses to see if there is an epistatic interaction and worsening of the neural crest related phenotype could give more information.

6.4.2.5 Sprouty4 (Spry4) and Syndecan2 are upregulated in Sp^{2H}/Sp^{2H}

Spry4 is a novel receptor tyrosine kinase pathway antagonist that has been shown to inhibit angiogenesis in mouse embryos (Lee et al. 2001). Its expression levels are increased by 40% in *Pax3* null embryos. *In vitro* studies have shown that *Spry4* arrests cell cycle progression and inhibits cell migration by inhibiting VEGF and FGF signalling pathways (Lee et al. 2001).

Syndecan2 is a cell surface heparin sulphate proteoglycan that is essential for angiogenic sprouting during embryogenesis in zebrafish (Chen et al. 2004). *Vegf* and *syndecan2* genetically interact *in vivo* in zebrafish (Chen et al. 2004). A yeast two hybrid screen using syndecan2 as bait, identified Syndecan Binding Protein (SDCBP) as a binding partner (Grootjans et al. 1997). A second two hybrid screen identified Sox4 as a binding partner of SDCBP (Geijsen et al. 2001) and luciferase assays showed that SDCBP acts as an adaptor molecule in interleukin-5 receptor-α-subunit-mediated activation of Sox4 (Geijsen et al. 2001). Homozygous *Sox4* null embryos have PTA and other OFT defects (Schilham et al. 1996). From this microarray experiment *syndecan2* appears to be significantly upregulated in Sp^{2H}/Sp^{2H} (1.5 fold) and $Sp^{2H}/+$ (1.6 fold) embryos. It is not yet known whether Syndecan2 binding to SDCBP has an effect on *Sox4* activation, however this is a pathway that might be relevant to the OFT defects in Sp^{2H}/Sp^{2H} embryos.

6.4.2.6 Aggrecan is downregulated in Splotch embryos

The cardiac defects in *splotch* embryos are due to defects intrinsic to the neural crest cells themselves as well as the extracellular environment through which they migrate (Chan et al. 2004). Marked overexpression of transcripts of *versican* in the pathways of neural crest migration, suggest that overexpression of this molecule leads to arrest of NCC migration in *splotch* embryos (Henderson et al. 1997).

Aggrecan is a chondroitin sulphate proteoglycan similar to *versican* but with a different expression pattern. *Aggrecan* is downregulated to 40% and 20% of wild type levels in Sp^{2H}/Sp^{2H} and $Sp^{2H}/+$ embryos respectively and this is confirmed by RTQPCR. *Aggrecan* is expressed in the peri-notochordal extracellular matrix where its normal function is to repel NCCs away from this region and towards their ultimate goal in the OFT. It is possible that reduced *aggrecan* expression disrupts the repulsion which may result in subtle changes in the NCC numbers and timing of migration from the neural tube, leading to the defects seen in Sp^{2H}/Sp^{2H} embryos. Alternatively the delay in migration of NCCs seen in Sp^{2H}/Sp^{2H}, could cause a downregulation of *aggrecan*, to try and reduce the inhibitory factors in the extracellular environment and encourage NCC migration. However, the expression pattern of *aggrecan* is not changed in Sp^{2H}/Sp^{2H} embryos, so if indeed there is downregultion, the changes are too subtle to see by *in situ* hybridisation (Henderson et al. 1997).

6.4.3 Neurogenesis related genes

The peripheral nervous system consists of multiple neural lineages derived from the neural crest. *Pax3* has been shown to regulate the differentiation of peripheral neurons (Koblar et al. 1999). In addition the *Sp/Sp* mutant is known to have abnormalities within the cranial ganglia and nerves (Tremblay et al. 1995) and *Sp/Sp* have loss of dorsal root and sympathetic ganglia (Auerbach 1954). A peripheral neuropathy similar to Charcot-Marie-Tooth type 1 is experienced by some individuals with WS4 caused by mutations with *SOX10* which is known to physically interact with *PAX3* (Inoue et al. 1999; Bondurand et al. 2000). With these facts in mind it is interesting to examine the neurogenesis related differentially expressed genes from this microarray experiment.

6.4.3.1 Hmx1 is significantly decreased in Sp^{2H}/Sp^{2H}

Hmx1 is a homeobox gene expressed in the dorsal root, sympathetic ganglia and in the trigeminal (V) and facio-acoustic ganglia (VII and VIII). At E10.5 it is also expressed in the 1st and 2nd branchial arches {Wang et al. 2000}. *Hmx1* is decreased in Sp^{2H}/Sp^{2H} embryos and this is confirmed by RTQPCR. It is possible that this apparent "downregulation" is due to tissue loss as Sp^{2H}/Sp^{2H} embryos lack dorsal root and cranial ganglia. However the demonstration that *Hmx1* is also decreased in heterozygous embryos that have no tissue loss (by RTQPCR although not by microarray) hints at a possible role in the neurogenesis defects seen in Sp^{2H}/Sp^{2H} embryos.

6.4.3.2 Myelin basic protein (MBP) is upregulated in Sp^{2H}/Sp^{2H} embryos

MBP is a known transcriptional downregulated target of *Pax3* and *Pax3* has been shown to bind to a sequence on the *MBP* promoter and cause transcriptional repression *in vitro* (Kioussi et al. 1995). This was corroborated by its upregulation in this microarray experiment to 70% above wildtype in Sp^{2H}/Sp^{2H} embryos. Some NCCs ultimately differentiate into myelin Schwann cells and form the myelin sheaths surrounding neurons. *Pax3* dysfunction is associated with ablation of the Schwann cell lineage (Epstein et al. 1994). It is possible that in wild type embryos *Pax3* functions to maintain NCCs in a precursor state and this involves suppression of *MBP*. It is postulated that at later stages in mouse development (E18 and beyond), *Pax3* becomes downregulated to allow *MBP* to upregulate and produce the myelin sheath of Schwann cells (Kioussi et al. 1995).

6.4.3.3 Two genes linked to Charcot-Marie-Tooth disease are downregulated in Sp^{2H}/Sp^{2H} embryos

Kinesin *Kif1B* is a gene found to be mutated in human Charcot-Marie-Tooth type 2A (Zhao et al. 2001). In this disease patients suffer from a peripheral neuropathy, muscle weakness and atrophy. *Kif1B* is responsible for microtubule mediated axonal anterograde transport in neurons and it was found to be significantly decreased to 20% of normal levels in Sp^{2H}/Sp^{2H} embryos in my microarray study.

Calnexin is a gene found to be decreased in these microarray data, although its relative change was only 0.71 in homozygote Sp^{2H}/Sp^{2H} embryos and 0.58 in heterozygote embryos so it was not included in table 6.3.3.1. It is a ubiquitously expressed membrane protein that is exclusively localised to the endoplasmic reticulum (ER) where it acts as a chaperone. *Calnexin* has been shown to interact with *PMP22* which is a Schwann cell derived peripheral myelin protein found to be mutated in Charcot-Marie-Tooth disease (type 1).

Both *Kif1B* and *calnexin* are expressed at decreased levels in Sp^{2H}/Sp^{2H} embryos. At this stage it is unclear whether this is due to reduced neuronal cell numbers in Sp^{2H}/Sp^{2H} or whether they have a more direct role in mediating the abnormal neurogenesis seen in Sp^{2H}/Sp^{2H} embryos.

6.4.3.4 Neural Proliferation and Differentiation Control gene 1 (NPDC1) is significantly decreased in Sp^{2H}/Sp^{2H} embryos and its transcriptional target E2F-1 is increased

From experiments by Koblar *et al* it has been suggested that *Pax3* has a role in differentiation of peripheral neurons (Koblar et al. 1999). However, genes that might mediate this differentiation by *Pax3* have not yet been identified. *NPDC1* is expressed at 30% wild type levels in Sp^{2H}/Sp^{2H} embryos. It is known to decrease cell proliferation and is expressed when neuronal cells undergo their terminal differentiation (Galiana et al. 1995). *NPDC1* has also been shown to bind to *E2F-1* and reduce its transcriptional activity (Dupont et al. 1998; Sansal et al. 2000). *E2F-1* is increased in this microarray study by 1.6 fold in Sp^{2H}/Sp^{2H} embryos. It is possible that changes in these two genes could mediate the reduced neuronal differentiation seen in Sp^{2H}/Sp^{2H} embryos.

6.4.3.5 Other Genes involved in neurogenesis

Several other genes involved with the development of the nervous system are significantly changed in Sp^{2H}/Sp^{2H} embryos, but their significance is uncertain. Table 6.4 shows some of these changed genes. For instance GCS is an enzyme that catalyzes the formation of glucosylceramide from ceramide, providing a route for ceramide clearance. The products of *GCS* enzymatic activity are gangliosides which are essential for differentiation and migration of cells in the nervous system (Buccoliero et al. 2002). In Sp^{2H}/Sp^{2H} embryos GCS is decreased to 30% of wild type levels in E10.5 embryos

from this experiment. It is plausible that *Pax3* upregulates *GCS* in neural precursor cells to allow their differentiation into peripheral neurons. Alternatively it may be that this apparent "downregulation" is due to loss of *GCS* expressing dorsal root ganglion cells in Sp^{2H}/Sp^{2H} embryos.

Ephexin (a gene involved with axon guidance) is decreased to 30% and 60% of wild type levels in homozygote and heterozygote Sp^{2H} embryos respectively. *Neurogenic differentiation 2* and *Mafk* (both involved with neural differentiation in the central nervous system), *astrotactin* (important for neuronal migration) and *neurocan* are all increased in Sp^{2H}/Sp^{2H} to between 1.5 and 2.2 fold in this microarray experiment. *Neurocan* in particular is interesting as it has been shown to bind to *N-CAM* with high affinity (Friedlander et al. 1994), a gene that it is activated by *Pax3* (Edelman & Jones 1998), although the significance of this interaction is unclear.

6.4.4 Genes involved with Neural tube Closure

One of the major structural defects seen in Sp^{2H}/Sp^{2H} embryos is an abnormality of neural tube closure (exencephaly and spina bifida) (Auerbach 1954; Beechey 1986; Franz 1992). Although the main purpose of this microarray experiment was to ascertain developmental pathways pertaining to the neural crest defects apparent in Sp^{2H}/Sp^{2H} embryos, a segment of the neural tube was included in the branchial arch dissections and a few genes related to the neural tube defects (NTDs) were uncovered.

Hic-1 is a gene with tumour suppressor activity that is included within the deletion in Miller-Dieker syndrome (Grimm et al. 1999). *Hic-1* deficient mice develop exencephaly and other severe developmental defects (Carter et al. 2000). *Hic-1* is significantly increased to levels of 1.9 fold in Sp^{2H}/Sp^{2H} embryos. It would be interesting to investigate whether transheterozygotes from *hic-1* deficient and *Pax3* deficient mice matings rescue the NTD phenotype seen in these embryos.

RXRα is upregulated in Sp^{2H} nulls to 1.8 fold of wild type levels in this experiment. Homozygous *RXRα* embryos also suffer from NTDs (Gruber et al. 1996) indicating a possible feedback upegulation of this gene in Sp^{2H}/Sp^{2H} embryos to compensate for absent *Pax3* .

6.4.5 Regulators of Somite development

Pax3 is a key gene involved with somite development. In Sp^{2H}/Sp^{2H} embryos tissues of somite origin are severely affected. During somitogenesis *Pax3* expression occurs in the unsegmented paraxial mesoderm, throughout the newly formed epithelial somites and then becomes localised to the dermatomyotome (Williams & Ordahl 1994). Later *Pax3* is expressed in the ventrolateral region of the dermatomyotome from which hypaxial limb muscles arise. In Sp^{2H}/Sp^{2H} embryos absence of functional *Pax3* causes failure of migration of myoblasts into the limb buds leading to a reduced muscle mass. The intercostal and body wall muscles, derivatives of the ventrolateral myotome, are also abnormal in Sp^{2H}/Sp^{2H} homozygotes (Franz 1993; Tajbakhsh et al. 1997).

Pax3 mutations also lead to malformation of the vertebral column and ribs which develop from the somitic sclerotome (Tremblay et al. 1998; Henderson et al. 1999). *Pax3* is not expressed in the sclerotome, however it is expressed in the paraxial mesoderm that gives rise to the somites (Schubert et al. 2001). Sp^{2H} homozygotes have a loss of spatial somite organisation perhaps resulting in a loss of signalling between somitic departments (Henderson et al. 1999). Several genes related to somite development are differentially expressed in Sp^{2H}/Sp^{2H} embryos in this microarray experiment (see figure 6.4) including lunatic fringe, *Wnt6, Notch1* and *scleraxis.*

6.4.5.1 Lunatic Fringe is decreased in Sp^{2H}/Sp^{2H} embryos

Lunatic Fringe (LF) is a secreted protein that potentiates *Delta* dependent *Notch-1* activation (Panin et al. 1997) and inhibits *Serrate/Jagged* dependent *Notch-1* activation in Drosophila (Fleming et al. 1997). Recently autosomal recessive mutations have been identified in LF in individuals affected by spondylocostal dysostosis which causes severe vertebral anomalies (OMIM #609813; Sparrow et al. 2006). *LF*'s expression is controlled in waves as it contributes to the timed formation of somites (Serth et al. 2003). Its expression has previously been examined in Sp^{2H}/Sp^{2H} and wild type embryos at E10.5 in the thoracic region where it is expressed in the somitic borders (Schubert et al. 2001). This pattern as found to be disturbed in Sp^{2H}/Sp^{2H} where *LF* signals were irregular or missing (Schubert et al. 2001). Its significant downregulation in Sp^{2H}/Sp^{2H} and $Sp^{2H}/+$ embryos in this microarray experiment and confirmation by RTQPCR was therefore able to serve as a positive control of a gene that was known to be decreased.

Interestingly it has been shown that the expression of the *Notch-1* ligand *Delta* is disrupted in Sp^{2H}/Sp^{2H} embryos (Schubert et al. 2001). As *LF* potentiates *Delta* dependent *Notch-1* activation, the downregulation of *LF* seen in Sp^{2H}/Sp^{2H} could result in loss of *Delta*'s control. Interestingly *Notch-1* was found to be upregulated by two fold in Sp^{2H}/Sp^{2H} embryos perhaps to compensate for abnormal *Delta* regulation, although the raw signal intensity levels were not robust enough to pass the statistical filter applied.

6.4.5.2 Other somitic genes are differentially expressed in Splotch embryos

The sclerotome and myotome regions of the somites are substantially disorganised in Sp^{2H}/Sp^{2H} embryos (Schubert et al. 2001). It is therefore noteworthy that *scleraxis* (a marker for lateral sclerotome) and *calsequestin* (a myotome marker) are significantly increased in this microarray experiment by 1.9 and 1.4 fold respectively. However the expression pattern for *scleraxis* in wild type and Sp^{2H}/Sp^{2H} embryos showed no obvious changes (Schubert et al. 2001).`

In addition to involvement in OFT formation, *Pkd1* is also expressed in the differentiating cartilage and somites and earlier in the notochord and floor plate which are involved in somite development. *Pkd1* deleted mice have disrupted vertebrae with curved and twisted spines and the abnormal development of long bones (Boulter et al. 2001).

Wnt6 is increased by 2.3 and 2.7 fold in Sp^{2H}/Sp^{2H} and $Sp^{2H}/+$ embryos respectively. Interestingly a recent study in chick embryos has shown that *Wnt6* regulates the epithelialisation process of somite formation (Schmidt et al. 2004). By overexpressing *Wnt6,* somites were maintained in their epithelialised state. The somites showed a delay in sclerotome differentiation and the number of migrating muscle precursors was reduced leading to the development of limbs with a reduced myogenic content. *Pax3* expression was found to be increased in chick embryos with increased *Wnt6* compared to controls (Schmidt et al. 2004). It is therefore plausible that reduced *Pax3* levels seen in *Splotch* homozygotes and heterozygotes in this microarray experiment have fedback to cause compensatory increased *Wnt6* levels.

6.4.6 Apoptotic pathway genes are affected in Splotch

Apoptosis is an important physiological process in development and has been demonstrated to occur in excess in some of the cellular pathways disrupted in *Splotch* embryos. In fact several observations support the notion that *Pax3* might have a physiological role in controlling apoptosis.

It has previously been demonstrated that in *Sp/Sp* embryos, neural tube defects are associated with neuroepithelial apoptosis (Phelan et al. 1997), suggesting that inhibition of apoptosis by *Pax3* might be an important feature for the survival of cells in embryogenesis. Apoptosis is also prevalent in the somites and dorsal root ganglia of *Splotch* embryos (Borycki et al. 1999). However apoptotic cardiac NCCs were not detected in abundance in the OFTs of Sp^{2H}/Sp^{2H} embryos, possibly because high levels of apoptosis in the endocardial cushions in wild type embryos makes comparative studies difficult (Epstein et al. 2000b).

Pax3 is well known for its involvement in the transformation of muscle cells in rhabdomyosarcomas (RMS). RMS is a soft tissue sarcoma with characteristic features of myogenic cells that presents mainly in paediatric patients. There are two main types of RMS: embryonal (eRMS) and alveolar (aRMS). eRMS often presents with mutations in the *p53* tumour suppressor gene and an aberrant expression of *PAX3* (Frascella et al. 1998). The more aggressive form of RMS, the alveolar subtype, is characterised by translocations between chromosomes 2 and 13 (t(2;13)(q35;q14)) resulting in chimeric genes encoding the fusion transcription factor *PAX3/FKHR*. The PAX3/FKHR fusion protein is a potent transactivator and is able to transform murine fibroblast cells in culture (Scheidler et al. 1996). Furthermore downregulation of both *PAX3/FKHR* in aRMS and *PAX3* in eRMS by antisense oligonucleotides revealed that the tumour cells underwent apoptosis (Bernasconi et al. 1996).

Evidence to date has demonstrated the involvement of *Pax3* with at least two genes involved with cell cycle regulation. There is direct transcriptional activation of the anti-apoptotic protein *BCL-XL* by *PAX3* and *PAX3/FKHR in vitro* (Margue et al. 2000). *Pax3* might also inhibit *p53* dependent apoptosis, as neural tube defects in *Splotch* embryos were rescued by *p53* loss of function (Pani et al. 2002).

6.4.6.1 Many genes involved with apoptotic control are differentially expressed in Splotch embryos

Figure 6.4 lists the differentially regulated apoptotic related genes that passed the stringent statistical analysis applied to the gene list.

6.4.6.1.1 Hic1 is implicated in the aetiology of RMS and is upregulated in Splotch

Hic1 or *hypermethylated in cancer-1* encodes a zinc finger transcription factor and its function in embryogenesis has already been discussed above. Disruption of one germ line allele in mice has been shown to produce malignancy including a preponderance of sarcomas in female mice (Chen et al. 2003). Hypermethylation resulting in inactivity of *Hic1* has been demonstrated in 59% of RMS cell lines (100% of aRMS and 33% of eRMS) (Rathi et al. 2003). *Hic1* is increased two fold in Sp^{2H}/Sp^{2H} embryos making it an interesting candidate gene for repression by *Pax3*.

6.4.6.1.2 Differential expression of Stat genes and their interactors in Splotch

Other changed genes resulting in a net inhibition of apoptosis are *Stat-3* (increased 1.5 fold in Sp^{2H}/Sp^{2H} in this experiment) and *Stat-1* (expressed at 30% wild type levels in Sp^{2H}/Sp^{2H} in this experiment). *Stat-1* is able to inhibit *Bcl-2* and *Bcl-X* and *Stat-3* activates *Bcl-2* and *Bcl-X* resulting in an overall anti-apoptotic effect (Stephanou et al. 2000).

FancC has been shown to activate *Stat-1* (Pang et al. 2000) and *FancC* is increased two fold in Sp^{2H}/Sp^{2H}, perhaps as a compensatory mechanism. Conversely *Protein Inhibitor of Stat-3 (PIAS3)* is downregulated to 30% wild type levels and this downregulation has been confirmed in Sp^{2H}/Sp^{2H} by RTQPCR in this experiment. *PIAS3* inhibits *Stat-3* (Chung et al. 1997). It is possible that the increase in *Stat-3* is as a result of release of this inhibition in Sp^{2H}/Sp^{2H} embryos.

PIAS3 has been shown to inhibit transcription of *MITF in vitro* (Levy et al. 2002), which is of relevance to Waardenburg syndrome as WSII is caused by mutations in *MITF* (Tassabehji et al. 1994). Watanabe *et al* demonstrated that *Pax3* activates *MITF* transcription *in vitro* (Watanabe et al. 1998). One can postulate that *PIAS3* is downregulated to try and compensate for a reduction of *MITF* in *Splotch* embryos.

It is possible that if certain genes are differentially expressed causing a net pro-apoptotic response when *Pax3* is non functional, that the converse could be true when *Pax3* has increased transcriptional activity, as in the *PAX3/FKHR* fusion gene. This may have important implications for tumourigenesis.

6.5 Summary

To reduce the number of false positive and false negatives only genes that showed consistent changes and passed stringent statistical tests were considered as possible direct or indirect targets of *Pax3*. Clearly further work with quantification by RTQPCR and assessment of qualitative changes using *in situ* hybridisation is needed to add further evidence for specific changes.

Tissue loss artefact is a potential problem with the homozygous *Pax3* results. For instance *Hmx1* and *Hira* are expressed in NCCs, which are reduced in number in Sp^{2H}/Sp^{2H} embryos. Conversely, coexpression in the same tissue type as *Pax3* increases the likelihood that they are direct transcriptional targets. Either way, such genes add confidence to the methodology of microarray expression arrays to detect gene changes.

Verification in the form of bioinformatic techniques would provide some information on promoter regions of potential target genes. However ultimately *in vitro* promoter assays are needed to define direct targets. Generation of murine transheterozygotes for phenotype assessment may provide definitive evidence for involvement of both genes in related pathways.

Some of the genes identified are likely to be direct or indirect target of *Pax3*. It is difficult to postulate which expression changes described above are primary or compensatory without further experimental work. The identification of some known targets of *Pax3* adds support to the ability of this methodology to detect biologically relevant changes.

Chapter 7 Investigation of patients by array-CGH referred as suspected cases of 22q11DS

7.1 Introduction

The majority of patients affected by the 22q11DS phenotype have a 3Mb deletion on chromosome 22q11.2 (Carey et al. 1992) which is detectable using FISH with the *TUPLE1/HIRA(D22S75)* probe. However there are many patients with features of 22q11DS without a known cause justified by conventional karyotype, FISH analysis or mutation screens for *TBX1* (Gong et al. 2001);(Yagi et al. 2003). Recent advances in molecular cytogenetic technology have allowed analysis of these patients by using microarray-comparative genomic hybridisation (array-CGH) (Solinas-Toldo et al. 1997; Pinkel et al. 1998). Before this technique is considered, there are a number of other options available to the clinician which will be now reviewed.

7.1.1 Approach to non-deleted patients with a 22q11DS phenotype

Figure 7.1.1 illustrates a suggested approach to investigating patients presenting with the 22q11DS phenotype, but not deleted for the *TUPLE1* FISH probe.

7.1.1.1 Detection of patients with suspected atypical 22q11.2 deletions

One possibility is that patients not deleted for the *TUPLE1* FISH, have an atypical deletion on 22q11.2 (see figure 1.4.2.3). Estimates of the prevalence of such deletions are in the order of 4-5% (Saitta et al. 2004; Rauch et al. 2005). The relatively new technique of Multiplex Ligation dependent Probe Amplification (MLPA) is available on a service basis at some British molecular laboratories for the detection of DNA copy number changes (Schouten et al. 2002). The MRC-Holland MLPA kit for DiGeorge syndrome (http://www.mlpa.com/p023.htm) provides detection of atypical 22q11.2 deletions together with deletions at 8 other loci harbouring deletions/genes linked to the 22q11DS phenotype. These are detailed in the table 7.1.1.

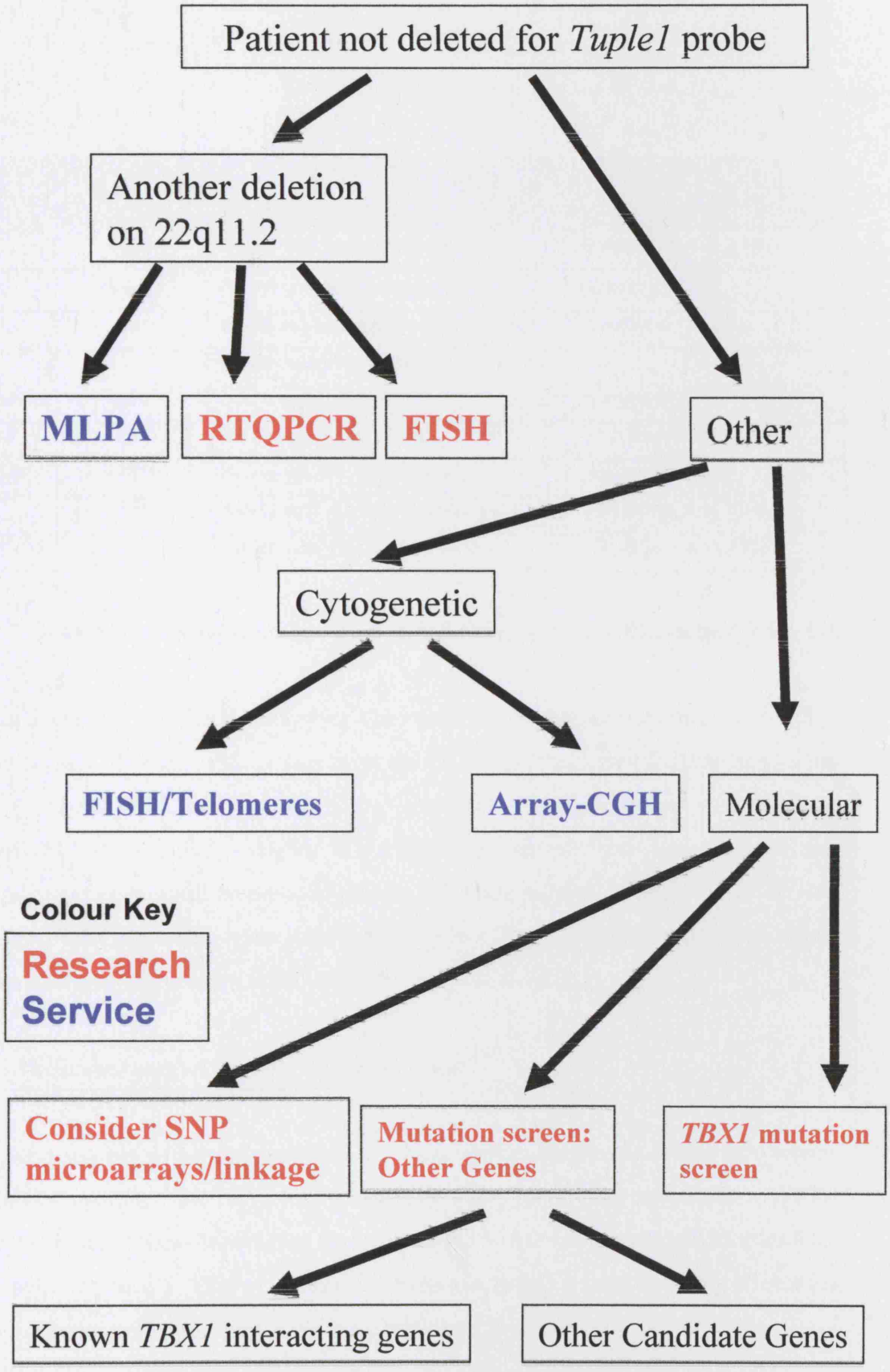


Figure 7.1.1

Suggested approach to investigating patients with the 22q11DS phenotype not deleted for the *HIRA* probe.

MLPA Probe	Chromosomal Position	Reason for Using Probe	Reference
NFKB1	4q24	Patient with phenotypic overlap.	(Fukushima et al. 1992)
CASP3	4q34.2	Patient with phenotypic overlap.	(Tsai et al. 1999)
HOXA3	7p15	Knock out mouse phenotypic overlap.	(Manley & Capecchi 1995)
GATA4	8p23	Mutations in congenital heart disease. Interacts with *TBX5*.	(Garg et al. 2003)
GATA3	10p15	HDR phenotypic overlap.	(Van Esch et al. 2000)
CUGBP2	10p14	Phenotypic overlap (DGCRII).	(Daw et al. 1996)
CRK	17p13	Patient with phenotypic overlap.	(Greenberg et al. 1988)
SERPINB7	18q21	Patient with phenotypic overlap.	(Greenberg et al. 1988)

Table 7.1.1 Details of MLPA probes used in the MRC-Holland DiGeorge MLPA kit.

Other reliable methods for investigating the possibility of atypical deletions in patients include the use of other FISH probes from the DiGeorge Syndrome Critical Region on 22q11.2 (Rauch et al. 2005), RTQPCR (Kariyazono et al. 2001) and microsatellite analysis (Shi et al. 2003). MLPA, RTQPCR and microsatellite analysis have the advantage that only small amounts of patient DNA are needed. MLPA, RTQPCR and interphase FISH have also been employed to detect 22q11.2 microduplications which produce a phenotypic overlap with 22q11DS (Yobb et al. 2005).

7.1.1.2 Molecular analysis of non deleted patients

Since the discovery of human mutations in *TBX1*, another diagnostic avenue is to screen non deleted patients for *TBX1* mutations. The paper describing mutations in *TBX1* emphasised that patients harbouring these mutations had fewer learning difficulties than 22q11 deleted patients (Yagi et al. 2003). Mutation analysis is only currently offered on a research basis as the prevalence of *TBX1* mutations is estimated to be quite low.

Patients where recessive inheritance is suspected (where there is parental consanguinity or more than one affected sibling and unaffected parents) can be considered on a research basis for homozygosity mapping studies via the use of standard linkage

techniques or application of SNP microarrays (Woods et al. 2004). Linkage analysis led to the discovery of mutations in *NKX2.6* in a consanguineous family with children affected with persistent truncus arteriosus (Heathcote et al. 2005).

Figure 7.1.1.2 illustrates examples of human genes found to be mutated in various syndromic and non-syndromic patients and murine models with phenotypic overlap with 22q11DS. Mutation analysis of some of these genes is offered on a service basis (*), but most are only available on a research basis. Patients with CHARGE syndrome (see table 1.2.2) have some features in common with the 22q11DS such as heart defects, developmental delay and ear anomalies. Mutations in *CHD7* (Vissers et al. 2004) and *SEMA3E* (Lalani et al. 2004) have been identified in CHARGE patients. Patients with Alagille syndrome (*JAG1 or NOTCH2*) or Noonan syndrome (*PTPN11*) can sometimes show phenotypic overlap with 22q11DS in that they have autosomal dominant craniofacial-heart syndromes.

Most of the other genes described in table 7.1.1.2 have been targeted for deletion in mouse models which show neural crest related abnormalities. Although some of these genes have been screened in human patients, it is currently impractical to screen every patient with a 22q11DS phenocopy. With ever improving technology in the form of sequencing microarray chips, this option might become available in the future.

7.1.1.3 Cytogenetic approach to non deleted patients

If patients referred as suspected cases of 22q11DS have had a normal karyotype and FISH using the *TUPLE1* probe, then there are a number of other cytogenetic options available. The first is FISH using probes for other loci. For instance some service labs have the FISH probe for the *GATA3* gene located at the HDR locus (Van Esch et al. 2000). Despite the fact that these patients have facial features that are very similar to 22q11DS patients, studies have shown that HDR is a comparatively rare syndrome and pursuit of deletions using this probe is minimally rewarding (Peter Scambler: personal communication). Hybridisation of probes for other loci such as *JAG1* (mutated in patients with Alagille syndrome) can also be considered if clinically suspected.

OFT: outflow tract
LD: learning difficulties
PTH: parathyroid hormone

** diagnostic mutation analysis available.*

Gene	Heart Defects			Defects							Other Defects	OMIM No
Syndromic	Aortic arch	OFT	General	LD	Craniofacial	PTH	Thymus	Renal	Ear Defects	Palate		
TBX1	✓	✓			✓	✓	✓	✓	✓	✓	Psychiatric	602054
*CHD7**			✓	✓	✓				✓	✓	Coloboma/Choanal Atresia	608892
SEMA3E			✓	✓	✓				✓		Coloboma/Choanal Atresia	603962
*PAX3/Pax3**	✓	✓			✓		✓		✓		Pigmentory/NTD	606597
*JAG1**		✓	✓	✓	✓			✓			Liver	601920
*PTPN11**		✓	✓	✓	✓				✓			176876
*TBX5**			✓								Radial Limb	601620
GATA3				✓	✓	✓		✓	✓			131320
CX43		✓			✓		✓		✓			121014
Pkd1		✓	✓					✓			Liver cysts	601313
HoxA3			✓			✓	✓				Thyroid hypoplasia	142954
Crkol	✓	✓			✓	✓	✓					602007
Cited 2	✓	✓							✓		Exencephaly	602937
Edn1	✓	✓			✓		✓					131240
Ece-1	✓	✓			✓		✓				Hirchsprung/autonomic	600423
Ednra	✓	✓			✓		✓		✓			131243
Edr1		✓			✓	✓	✓			✓	Spleen hypoplasia	602978
Fgf8	✓	✓			✓		✓		✓			600483
RaldH2	✓	✓			✓	✓					Thyroid hypoplasia	603687
Hand2	✓				✓							602407
Gbx2	✓				✓				✓	✓	Cerebellar defects	601135
Chordin	✓	✓	laterality			✓	✓		✓	✓	Hypoplastic thyroid, skeletal.	603475
Rxrα		✓	✓								Neural tune defects	180245
Non Syndromic												**OMIM No**
NKX2.6		✓										Heathcote *et al*
NKX2.5			✓									600584
GATA4			✓									600576
SEMA3C	✓	✓										602645
CFC1		✓										605194
PlexinD1	✓	✓										604282
BmpR1a		✓	✓									601299
Bmp4		✓										112262
Sox4	Rare	✓										184430
Fog2		✓	✓									603693
Foxh1		✓										603621

Table 7.1.1.2. Examples of human (capitals) and mouse genes causing overlapping 22q11DS phenotypes.

Probes to the telomeres of chromosomes are available on a diagnostic basis to assess patients that are severely delayed with multiple congenital abnormalities. Such patients are sometimes found to harbour cytogenetically invisible deletions affecting the telomeric chromosomal ends, detectable by using a panel of telomeric FISH probes (Ravnan et al. 2005).

Array-CGH is a molecular cytogenetic approach that is being used increasingly as a service and an accessible research tool, to aid in the investigation of patients with suspected chromosomal anomalies, who have a normal karyotype.

7.1.2 Array-Comparative Genomic Hybridisation

Array-CGH is a powerful technology enabling detection of submicroscopic chromosome duplications and deletions by comparing a differentially labelled test sample to a normal control (figure 7.1.2). The samples are co-hybridised to a microarray containing genomic clones and the resulting ratio of fluorescence intensities on each array element is proportional to the DNA copy number difference. The 1Mb whole genome array has a resolution which is 5-10 times greater than that provided by conventional karyotype analysis. It is also more sensitive than CGH applied to chromosome metaphase spreads, which can typically detect changes in the order of 5-10Mb (Kallioniemi et al. 1992), or 3-5Mb at high resolution (Kirchhoff et al. 2001). The technique of array-CGH itself has a resolution limited only by the clone size and clone density on the arrays.

Array-CGH has been used extensively to establish copy number changes in tumour samples. This approach has been employed both with region specific arrays (Monni et al. 2001; Redon et al. 2002; Wessendorf et al. 2003) and with genome wide arrays (Pollack et al. 2002; Wilhelm et al. 2002; Veltman et al. 2003). Development of arrays for use in this field is driven by their importance in diagnostic, prognostic and research applications, all ultimately affecting management.

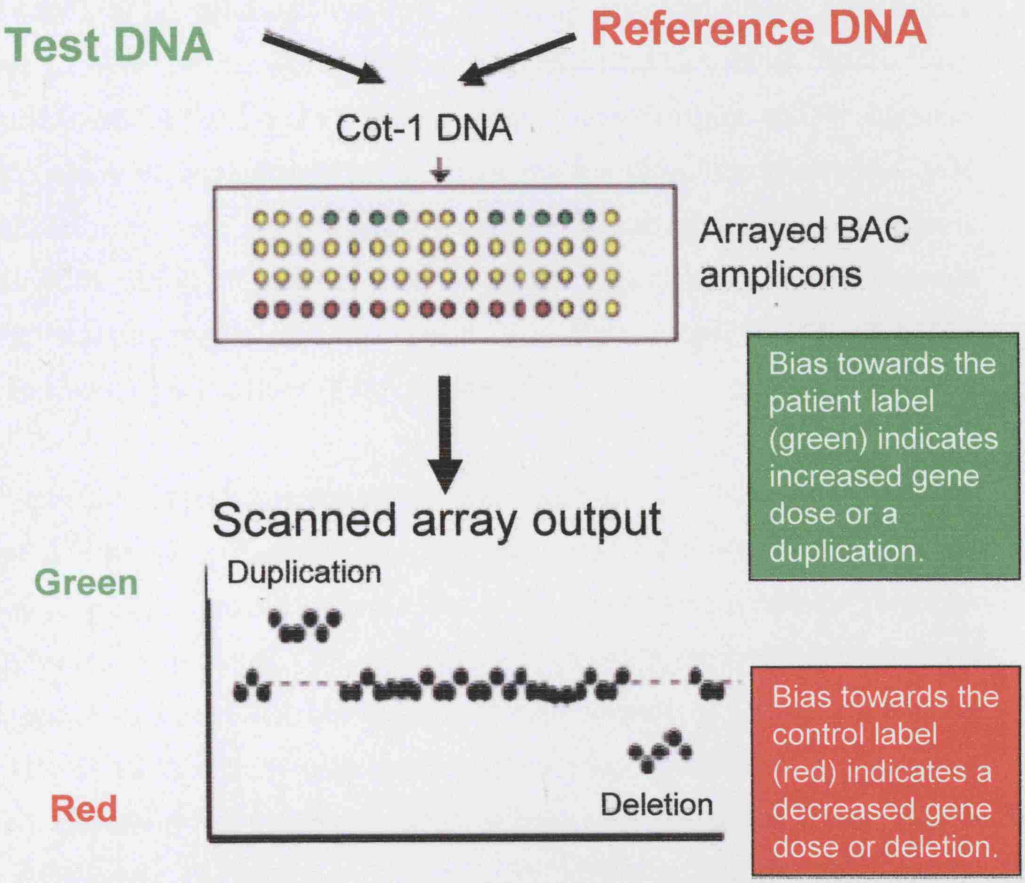


Figure 7.1.2
DNA labelling and hybridisation in array-CGH. Patient and control DNA are labelled with red and green fluorochromes respectively and co-hybridised with Cot1 DNA to the microarray and the relative signal intensity is measured. Bias towards the patient label (green) indicates increased gene dose or a duplication. Bias towards the control label (red) indicates a decreased gene dose or deletion.

Constitutional chromosomal abnormalities are becoming increasingly recognised as a major cause of developmental malformations, thus the application of array-CGH in this field is expanding. Custom arrays focusing on specific genomic regions for certain diseases have been developed and tested at high resolution for diagnosis and research. In 2002 Buckley *et al* published the first chromosome specific array representing chromosome 22 with an average resolution of 75Kb (Buckley et al. 2002). They demonstrated its use for tumour DNA copy number change detection and for diagnosis in 22q11DS. Yu *et al* have developed an array for the diagnosis of terminal 1p36 deletions which is one of the most commonly observed mental retardation syndromes in humans (Yu et al. 2003). In addition an array specific for detection of subtelomeric chromosome rearrangements was developed to screen patients with congenital anomalies and learning difficulties (Veltman et al. 2002).

It is now becoming increasingly recognised that many patients with an undiagnosed combination of learning difficulties, complex congenital anomalies and dysmorphic features may have chromosomal abnormalities undetectable by conventional karyotype analysis. High resolution array-CGH with whole genome coverage is being employed to establish chromosomal copy number changes in such patients. If patients are selected strictly to include those with features commonly seen in chromosomal abnormalities, the detection sensitivity of duplications and deletions can be in the order of 15%. Vissers *et al* detected 5 chromosomal changes in 20 patients studied of which two were definite polymorphisms giving a potential sensitivity of 15% (Vissers et al. 2003). Shaw-Smith *et al* demonstrated a detection rate of 14% in a cohort of 50 patients with learning difficulties and dysmorphism (Shaw-Smith et al. 2004).

	Patient	Heart	Face	Ca↓	Thymus	Palate	Development	Other
1	PS2	Tetralogy of Fallot/ARSA	Dysmorphic	✓	Absent	CLP	DNK	arrinencephaly
2	PS3	IAA/VSD/ASD	Dysmorphic	DNK	DNK	DNK	DNK	coloboma
3	PS4	VSD/PA	Dysmorphic	DNK	Absent	CLP	DD	micropenis/deafness
4	PS5	Tetralogy of Fallot	Dysmorphic	Normal	DNK	Bifid Uvula	DD	
5	PS6	Coarctation	Dysmorphic	✓	DNK	DNK	DNK	
6	PS7	IAA	Dysmorphic	✓	Absent	DNK	DNK	
7	PS8	VSD/PA/MAPCAs	Dysmorphic	DNK	DNK	VPI	DD	
8	PS13	Normal	Dysmorphic	DNK	DNK	VPI	DD	
9	PS15	IAA	Dysmorphic	DNK	DNK	DNK	DD	corpus callosum agenesis
10	PS17	Normal	Dysmorphic	DNK	DNK	Hypernasal speech	DD	behavioural problems
11	PS18	VSD	Dysmorphic	✓	Absent	DNK	DNK	hypospadius/fits
12	PS19	Tetralogy of Fallot	Dysmorphic	DNK	DNK	CLP	DNK	
13	PS21	VSD/PS/PTA	Dysmorphic	✓	Absent	DNK	DD	hemivertebrae
14	PS24	ASD/PDA	Dysmorphic	✓	Hypoplastic	DNK	DNK	retinal coloboma
15	PS25	IAA/VSD/PDA	Dysmorphic	DNK	DNK	DNK	DNK	"incomplete" gut

IAA: interrupted aortic arch
PDA: patent ductus arteriosus
ARSA: aberrant right subclavian artery
VSD: ventricular septal defect
MAPCAs: multiple aorto-pulmonary collateral arteries
PA: pulmonary atresia
Ca↓: hypocalcaemia
CLP: cleft lip and palate
DD: developmental delay
VPI: velopharyngeal insufficiency
DNK: details not known

Table 7.2.1
Clinical summary of the features overlapping with 22q11DS seen in 15 patients tested on the 1Mb genomic arrays.

7.2 Methodology

7.2.1 Selection of patients

27 patients were selected for DNA analysis based on their phenotypic similarity to 22q11DS. These patients were all originally seen by a clinical geneticist and investigated extensively for the aetiology of their disease. They all had at least two out of six features commonly found in 22q11DS. All 27 patients had a normal karyotype and a normal FISH analysis using the *TUPLE1* probe from the 22q11DS region. As the referring clinicians had a high index of suspicion for 22q11DS, with parental consent their DNA was sent to Prof P. Scambler at MMU, ICH for further genetic analysis. Of the 27 patient DNAs sent to the Sanger centre for analysis, 15 were selected for array-CGH. The other 12 were either not hybridised due to lack of resources, or the DNA was too impure for chip hybridisation. (table 7.2.1).

Clinical details of PS5 were obtained from his foster parents by me. Clinical details of PS15 were obtained by Dr Koen Devriendt as the patient was seen by him in Belgium. Ethical approval was obtained locally by our collaborators at the Sanger centre. Local regional ethical approval had been obtained by Prof. Scambler (Appendices 3.1, 3.2, 4.1 and 4.2). In addition patients where DNA was recently acquired, were written to requesting specific consent for array-CGH (appendices 5.1 and 5.2). Results were given back to the referring clinicians by letter (Appendix 6).

7.2.2 DNA microarrays

Array-CGH was performed at the Sanger centre by Dr Kathryn Woodfine, Dr Charles Shaw-Smith and Lisa Rickman under a collaboration arranged between Prof Nigel Carter, Prof Peter Scambler, Dr Charles Shaw-Smith and myself. Details of methodology can be found at http://www.sanger.ac.uk/HGP/methods/cytogenetics/. A brief outline of the microarray methodology is given below. Follow up work including FISH analysis for patient PS5 and microsatellite analysis for patients PS5 and PS16 was performed by me at MMU, ICH.

Microarrays were constructed as previously described (Fiegler et al. 2003). Sequencing clones from libraries held at the Wellcome Trust Sanger Institute were obtained from the published Golden Path of the human genome. For the 1Mb whole genome array, large insert clones were selected to be spaced at approximately 1Mb intervals across the human genome and were arrayed in duplicate. Clones were spotted in triplicate. Clone details can be obtained from the Ensembl genome browser www.ensembl.org/Homo_sapiens/cytoview. DNA from the clones was isolated, amplified by DOP-PCR followed by amino-linking PCR and spotted onto 3D-link activated slides (Motorola, Schaumburg, IL, USA) using a MicroGrid II arrayer (BioRobotics, Boston, MA, USA).

7.2.2.1 DNA labelling, hybridisation and analysis

Figure 7.1.2 illustrates the general principal of DNA labelling and hybridisation in array-CGH. Patient and control DNA were differentially labelled as described (Fiegler et al. 2003). Briefly 450ng of patient DNA was labelled with dCTP-Cy3 and 450ng of control DNA was labelled with dCTP-Cy5 using a Bioprime Labelling kit (Invitrogen, Carlsbad, CA). Control DNA was obtained from a pool of 20 males to minimise the effects of common polymorphisms in DNA copy number. The control DNA was tested on the arrays for polymorphisms, and a list was compiled for all clones which appeared to be polymorphic. After labelling, the DNA was purified using a G50 spin column (Amersham Pharmacia, Buckinghamshire, UK). The labelled DNA was co-hybridised to the array after pre-annealing with Cot1 DNA to suppress genomic repeats. Hybridisations were carried out as described (Fiegler et al. 2003). The arrays were scanned using an Axon 4000B scanner (Axon Instruments, Burlingame, CA), and images quantified using "Spot" software (Jain et al. 2002).

7.2.3 Microsatellite analysis

For full details of primer sequences and techniques please refer to chapter 2. FAM or TET labelled polymorphic microsatellite markers from the region of interest on 5q11.2 and 8q12 were identified and ordered from QIAgen. Microsatellite analysis was conducted on peripheral blood DNA for patient PS5 and PS16. PCR amplifications of microsatellite repeats were carried out in accordance with standard protocols using

AmpliTaq Gold. Samples were run on a MegaBACE 1000 and analysed using Genetic Profiler software.

7.2.4 FISH analysis

BAC clones were ordered from BACPAC recourses, (http://www.chori.org/bacpac/), and DNA was isolated using standard alkaline lysis techniques. The clone DNA was labelled with SpectrumRed-dUTP or SpectrumGreen-dUTP by nick translation (Vysis, Inc, Illinois, USA). The clones were hybridised to chromosome metaphase spreads from patient PS5 by a modification of standard methods (Pinkel et al. 1988). Chromosomes were counterstained with 4',6-diamidino-2-phenylindole (DAPI). Results were visualised using a Leica DMLB fluorescence microscope (Leica Microsystems Wetzlar GmbH) and images were captured using a MetaSystems ISIS workstation MetaSystems GmbH, Altlussheim).

7.3 Results

7.3.1 Array-CGH

Table 7.2.1 illustrates the clinical details of the 15 patients chosen for array-CGH. Of the 15 patient DNA's that were hybridised, three were found to have chromosome imbalances. These are illustrated below in table 7.3.1.

Pt	Heart	Face	Ca↓	Thymus	Palate	Development	Array-CGH
PS5	TOF	Dysmorphic	No	DNK	Bifid Uvula	DD	del 5q11
PS15	IAA	Dysmorphic	DNK	DNK	DNK	DD	del 8q12
PS25	IAA/VSD/ASD	Dysmorphic	DNK	DNK	DNK	DNK	del 5p/dup5q

Table 7.3.1 Results of DNA hybridisation to 1Mb array. See table 7.2.1 for abbreviations. Del: deletion, dup: duplication.

7.3.1.1 A relative loss of at least 5Mb was identified on DNA from patient PS5 on 5q11.2

Patient PS5 was found to have a relative loss of at least 5Mb of chromosomal material on chromosome 5q11.2 between the clones RP4-592P18 and RP11-412L4 (figure 7.3.1.1). The clones with the relative loss were not on the list of known polymorphic clones compiled from the 20 control and 50 patient DNAs tested at the Sanger centre.

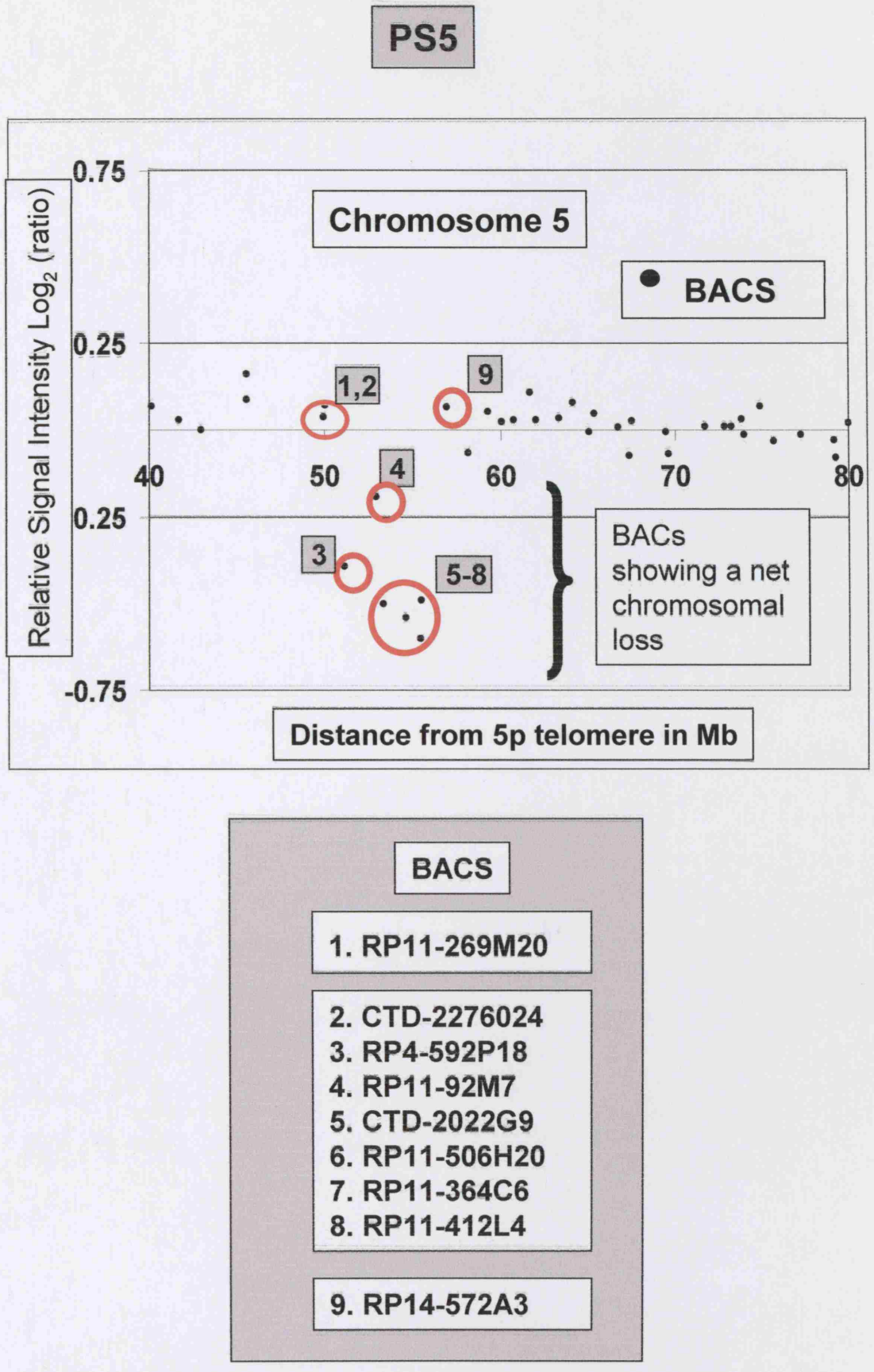


Figure 7.3.1.1

Microarray result for PS5 signifying a net chromosomal loss of at least 5Mb on 5q11. The log_2 of the measured fluorescence ratio is plotted versus the order of clones on chromosome 5p. Data are normalised so that log_2ratio=0 for portions of the genome that contain two copies of each sequence. Red circles: BACs labelled above.

7.3.1.1.1 Clinical details of patient PS5

DNA from patient PS5 was sent to Prof. Peter Scambler from Dr Maurice Super, Manchester in 1996 as his phenotype was noted to overlap with patients with 22q11DS and the patient was not deleted at 22q11. Patient PS5 was born at 37 weeks by emergency Caesarean Section for breech presentation and fetal distress. His birth weight was 2.25kg (2nd Centile). At 2 months of age he was admitted to hospital with lethargy and coffee ground vomiting and was found to be pale and shocked. After resuscitation he had left sided limb weakness and eye deviation. These neurological changes were thought to be secondary to hypoxic damage. An electroencephalogram and computerised tomography (CT) scan of his brain were normal. During the admission a pansystolic murmur was noted and tetralogy of Fallot was diagnosed by echocardiogram. He underwent repair of his heart anomaly at 18 months of age.

Horizontal nystagmus was noted from 6 months of age. The patient had surgery for a left divergent squint in 2001 and currently requires glasses for hypermetropia. From an early age he had at least three documented episodes of hypoglycaemic seizures and he currently suffers from absence seizures.

The patient had significant developmental delay. His speech was especially delayed and he first spoke in sentences from the age of 3 years. He has a bifid uvula and evidence of velopharyngeal insufficiency on videofluoroscopy testing. He has been in a school for children with learning difficulties from the age of 3 years. He had mild motor delay, first walking at 15 months and he currently suffers from motor dyspraxia and a tremor.

The patient also has significant behavioural difficulties. He has obsessional play and echolalia. He is said to be emotionally labile. There are particular problems with skin picking and a poor sleep pattern.

On examination at the age of 11 years the patient's height was 128.7cm (<3rd centile), weight 25.6kg (3rd-10th centiles) and head circumference was 50cm (< 3rd centile). His pictures are shown in figure 7.3.1.1.1. He had bilateral fetal fingerpads. He had a small mouth with irregular dentition, a bifid uvula and micrognathia. He had fixed flexion deformities of his knees. His feet showed a wide sandal gap with broad greater halluces.

Figure 7.3.1.1.1

Patient PS5 showing an asymmetrical face, prominent ears, thin lips and large broad greater halluces. Top images are PS5 at aged 4 years, bottom images are him aged 11 years.

Figure 7.3.1.1.2 Patient PS15 showing bitemporal narrowing, thin upper lip and low set overfolded ear helices.

Both pictures reproduced with consent from the families.

7.3.1.2 A relative loss of at least 6Mb was identified on DNA from patient PS15 on 8q12

Patient PS15 was found to have a relative loss of at least 6Mb of chromosomal material on chromosome 8q12 between the clones RP11-414L17 and RP11-115G12 (figure 7.3.1.2). The clones with the relative loss were not on the list of known polymorphic clones compiled from the 20 control and 50 patient DNAs tested at the Sanger centre.

7.3.1.2.1 Clinical details of patient PS15

PS15 is a female patient whose DNA was sent to Prof Peter Scambler from Dr Koen Devriendt, Leuven, Belgium. Her phenotype was noted to overlap with patients with 22q11DS, yet she had a normal karyotype and was not deleted for the *TUPLE1* FISH probe at 22q11.2. Figure 7.3.1.2.1 shows her facial features. She has severe learning difficulties with no speech development. She had major feeding difficulties during childhood requiring nasogastric tube feeding. An echocardiogram revealed that she had an interrupted aortic arch type A, preductal coarctation of the aorta, a large VSD, a 16mm secundum ASD and a bicuspid aortic valve. She was found to have an absent corpus callosum on CT. On examination she had short stature, long, slender fingers and low set overfolded helices.

7.3.1.3 Chromosome imbalance on PS25

PS25 was found to have a relative loss of 7.5-8.5Mb of chromosomal material from 5p15.31-pter and a relative gain of approximately 2Mb of material from 5q35.1-qter (figure 7.3.1.3). Similar chromosomal deletions on 5p have been described in patients with the cri du chat phenotype (OMIM 123450).

7.3.1.4.1 Clinical details of PS25

Patient PS25 was referred to Prof Peter Scambler by an American clinician as she had an overlapping phenotype with the 22q11DS, but no deletion detected by FISH. Clinical details are limited on this patient and an attempt to make further contact with the clinician following the abnormal array-CGH result failed. The only information

available on this patient was that she was born with coarctation of the aorta, aortic valve abnormalities, a VSD, a PDA, an "incomplete" gut, low set ears and mild facial dysmorphism. No photos were available.

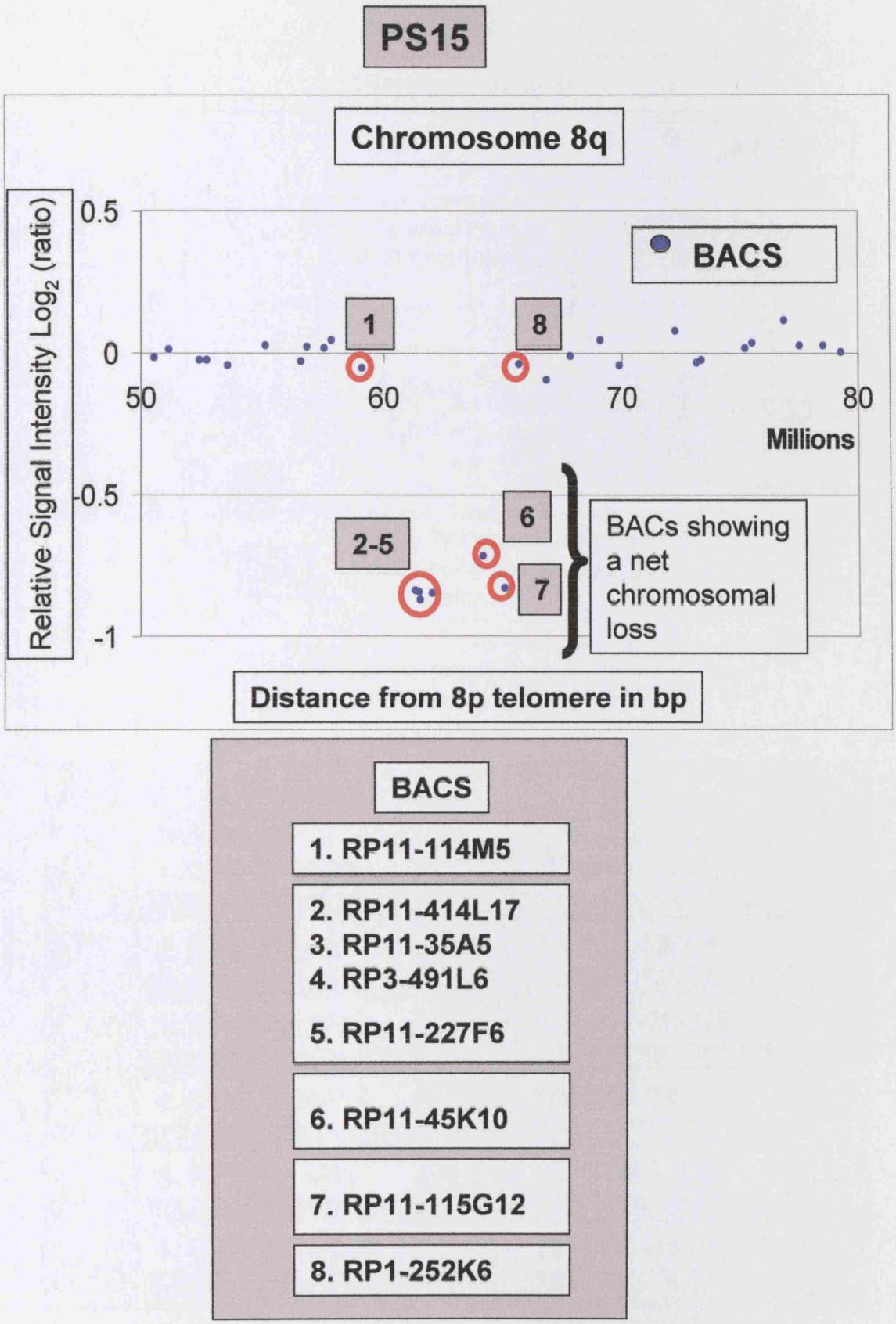


Figure 7.3.1.2

Microarray result for PS15 signifying a net chromosomal loss of 5-5.7 Mb on 8q12. The $\log_2$ of the measured fluorescence ratio is plotted versus the order of clones on chromosome 8p. Data are normalised so that $\log_2$ratio=0 for portions of the genome that contain two copies of each sequence. Red circles: BACs labelled above.

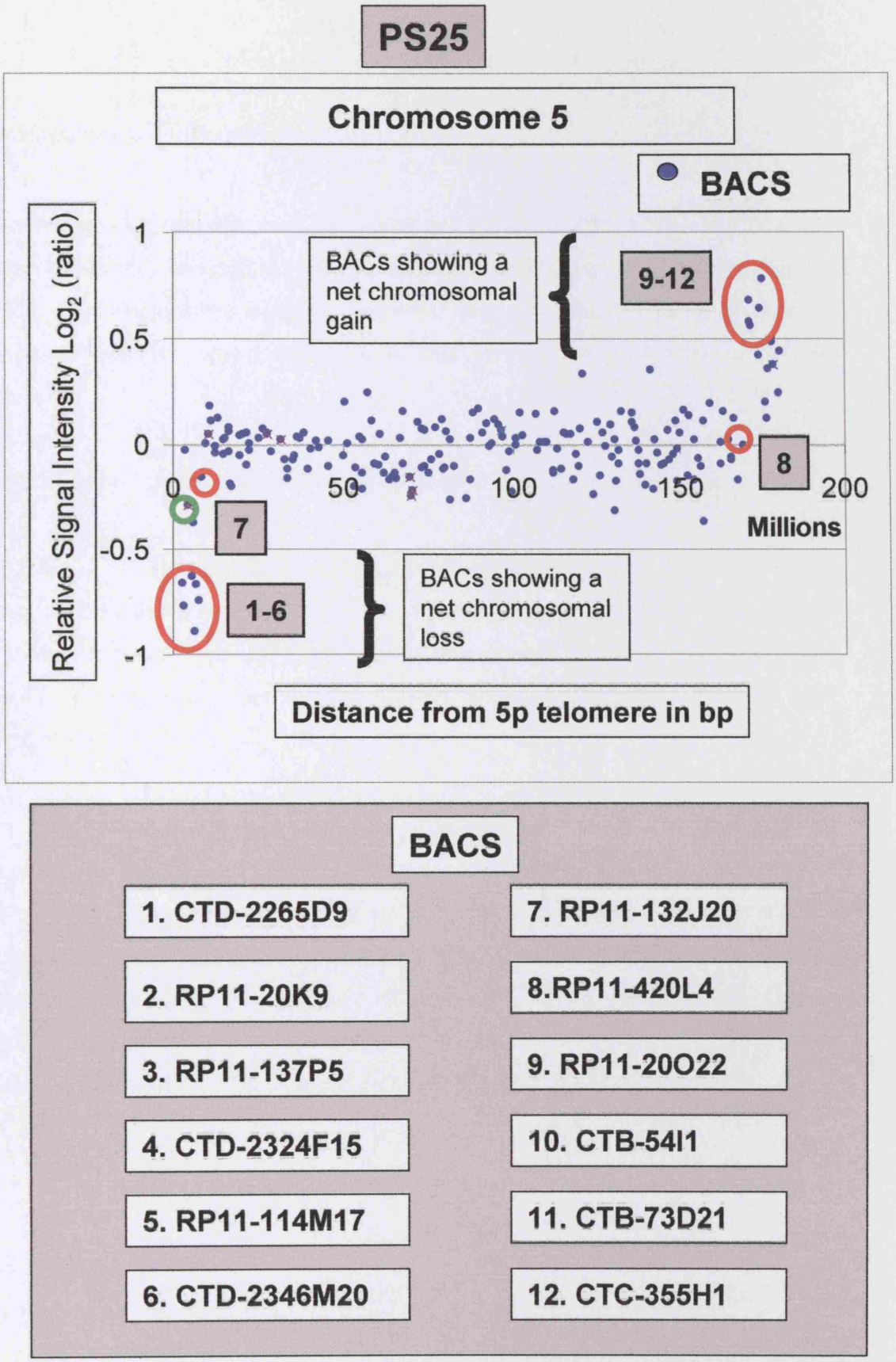


Figure 7.3.1.3

Microarray result for PS25 signifying deletion 5p15.3-pter (BACs1-6) and duplication 5q35.1-qter (BACs 9-12). The $\log_2$ of the measured fluorescence ratio is plotted versus the order of clones on chromosome 5p. Data are normalised so that $\log_2$ratio=0 for portions of the genome that contain two copies of each sequence. Red circles: BACs labelled above. Green circle: known problematic BAC clone RP11-13H10.

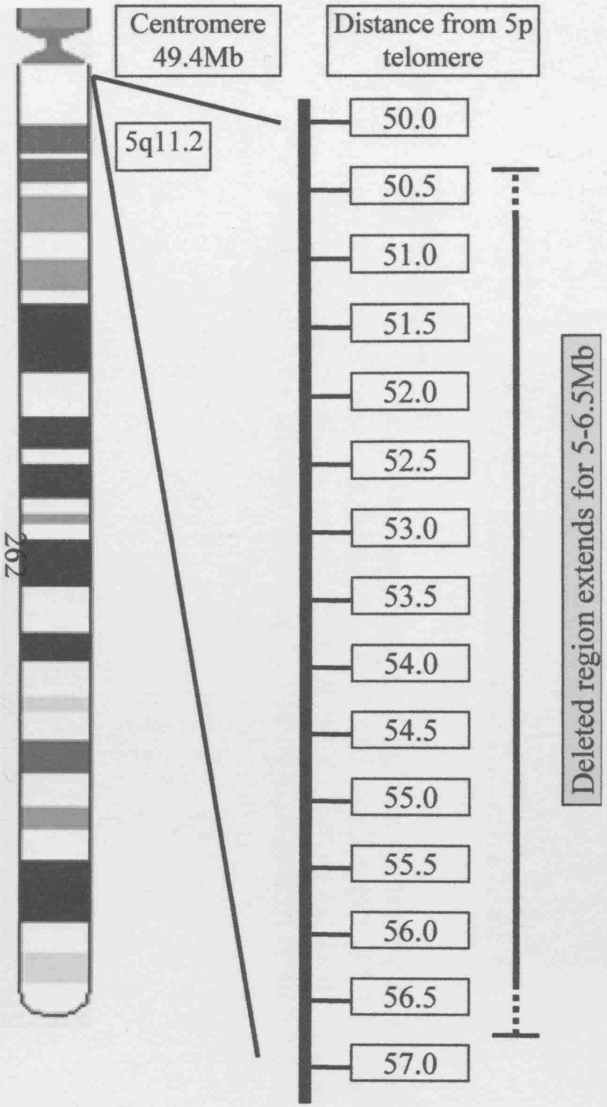
7.3.2 Validation of Results

FISH and microsatellite validation was performed by me on PS5. FISH validation was performed by Koen Devriendt on PS15. Microsatellite validation was performed by me on PS15. The chromosome imbalance on PS25 was not validated partly because no chromosomal material was available from this patient and partly due to lack of resources.

7.3.2.1 Validation of chromosomal loss on PS5 by FISH and microsatellite analysis.

Parental blood was unavailable for checking whether the occurrence of the deletion was *de novo,* as the child is in foster care with no parental contact. Despite two previously normal karyotypes, a repeat karyotype was performed and on closer inspection of the region it was possible to identify the deletion, although this would not have been detected routinely.

Figure 7.3.2.1 summarises the results of the validation experiments. To verify the array-CGH results for the BACs in the 5q11.2 region, FISH analysis was performed for six of the BACs confirming the results shown on the arrays (figure 7.3.2.1.1a-d). In addition microsatellite marker analysis was used to provide definitively dizygous flanking regions on the basis of heterozygosity of markers D5S466 and D5S2102. Due to absence of parental DNA and homozygosity of markers D5S491 and D21S264 it is possible that the deletion extends as far as 6.5Mb (figure 7.3.2.1).



Microsatellite Marker (chromosomal position)	Heterozygous/ Homozygous in Pt 5 (heterozygosity rate)	BAC/PAC clone (chromosomal position)	Retention/ Deletion of clone on array	Array result confirmed by FISH
		RP11-269M20 (50Mb)	Retained	Yes
		CTD-2276O24 (50Mb)	Retained	-
D5S466 (50.3Mb)	Heterozygous (0.49)			
D21S264 (50.5Mb)	Homozygous (0.75)	RP4-592P18 (51Mb)	Deleted	Yes
D5S2035 (51.3Mb)	Homozygous (0.53)			
D5S623 (52.3Mb)	Homozygous (0.78)			
D5S2037 (52.6Mb)	Homozygous (0.61)	RP11-92M7 (53Mb)	Deleted	Yes
		CTD-2022G9 (53Mb)	Deleted	-
D5S1969 (53.3Mb)	Homozygous (0.87)			
D5S1968 (53.6Mb)	Homozygous (0.77)			
D5S2076 (54.2Mb)	Homozygous (0.69)			
D5S664 (55.0Mb)	Homozygous (0.85)	RP11-506H20 (55Mb)	Deleted	Yes
		RP11-364C6 (55Mb)	Deleted	-
D5S2068 (55.6Mb)	Homozygous (0.81)			
D5S461(55.9Mb)	Homozygous (0.76)			
D5S407 (56.0Mb)	Homozygous (0.86)			
D5S491 (56.4Mb)	Homozygous (0.57)	RP11-412L4 (56Mb)	Deleted	Yes
D5S2102 (56.8Mb)	Heterozygous (0.75)			
D5S2507 (56.9Mb)	Heterozygous (0.8)	RP11-572A3 (57Mb)	Retained	Yes

Figure 7.3.2.1 Location of microsatellite markers and BAC/PAC clones on chromosome 5q11 used for identification and validation of the extent of the deleted region.

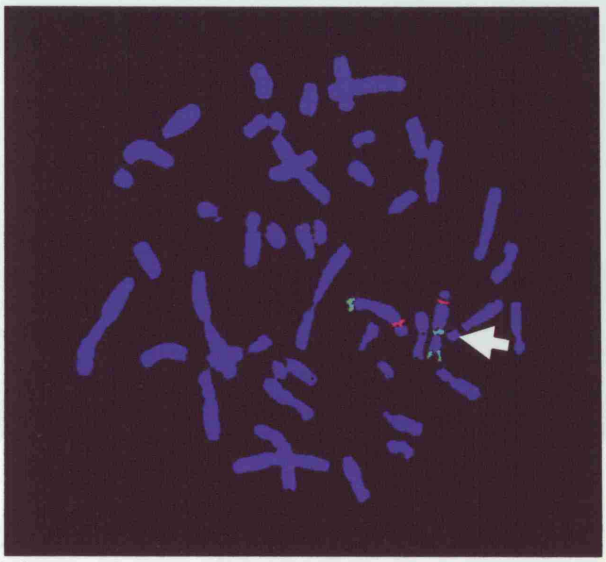


Figure 7.3.2.1.1 a
Fluorescence In-Situ Hybridisation to a metaphase spread of patient PS5 showing the absence of signal for BAC RP11-506H20 at 5q11.2 on one chromosome 5. White arrow: BAC RP11-506H20 signal (green); Distal green fluorescent signals: telomeric 5p Crit du Chat control probe; Red fluorescent signals: 5q EGR-1 control probe.

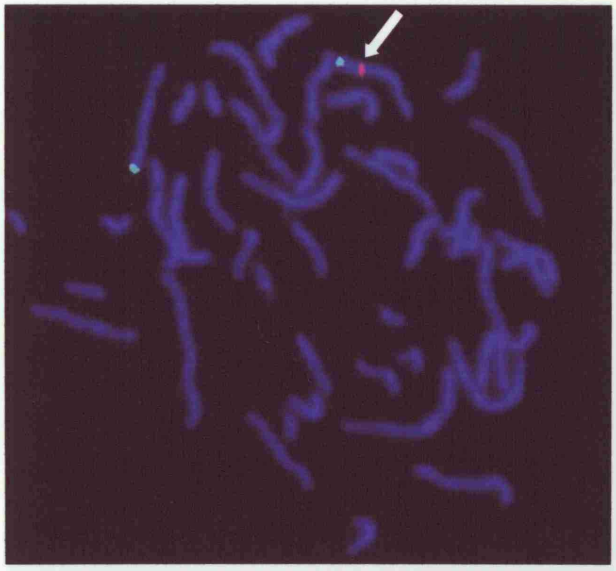


Figure 7.3.2.1.1 b
Fluorescence In-Situ Hybridisation to a metaphase spread of patient PS5 showing the absence of signal for BAC RP11-412L4 at 5q11.2 on one chromosome 5. White arrow: BAC RP11-412L4 signal (red); Distal green fluorescent signals: telomeric 5p Crit du Chat control probe.

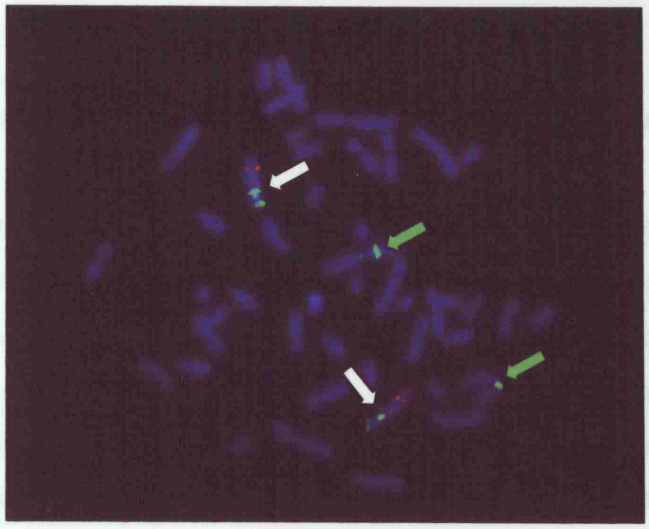


Figure 7.3.2.1.1 c
Fluorescence In-Situ Hybridisation to a metaphase spread of patient PS5 showing the presence of two signals for BAC RP11-269M20 at 5q11.2 on both chromosome 5s. White arrow: BAC RP11-269M20 signa (green)l; Distal green fluorescent signals: telomeric 5p Crit du Chat control probe; Red fluorescent signals: 5q EGR-1 control probe. Green arrows: cross hybridisation of BAC RP11-269M20 to chromosome 1q.

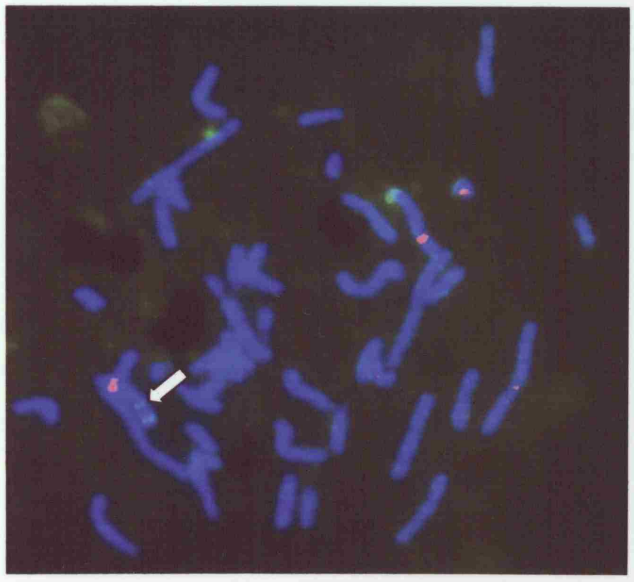
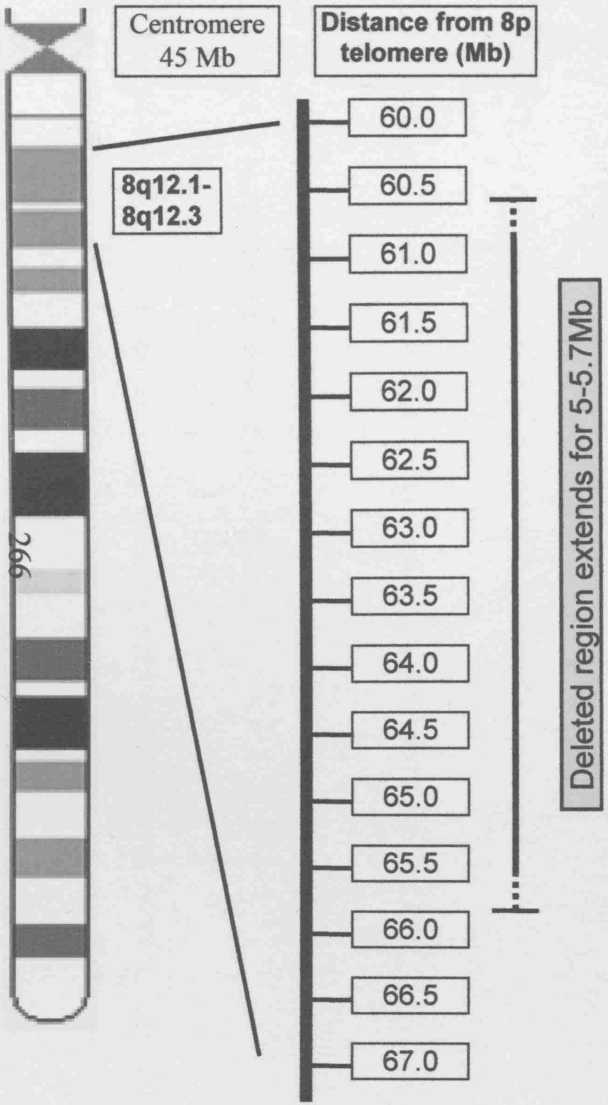


Figure 7.3.2.1.1 d
Fluorescence In-Situ Hybridisation to a metaphase spread of patient PS5 showing the absence of signal for BAC RP4-592P18 at 5q11.2 on one chromosome 5. White arrow: BAC RP4-592P18 signal (green); Distal green fluorescent signals: telomeric 5p Crit du Chat control probe; Red fluorescent signals: 5q EGR-1 control probe.

7.3.2.2 Validation of chromosomal loss on PS15 by FISH and microsatellite analysis.

Figure 7.3.2.2 summarises the results of the validation experiments that were performed for PS15. FISH and microsatellite analysis were used to confirm that the deletion and to identify deletion boundaries. Analysis of parental DNA by FISH (performed by Koen Devriendt) confirmed the deletion was *de novo*. Five labelled polymorphic microsatellite markers from 8q12 were used to amplify DNA from patient PS15 (heterozygosity rating for the marker in brackets). Heterozygosity of marker D8S1763 provided a proximal boundary for the deletion. The presence of BAC clone RP1-252K6 on both chromosome 8q12 regions provided a distal boundary for the deletion. The estimated size of the deletion is therefore 5-5.7Mb.



Microsatellite Marker (chromosomal position)	Heterozygous/ Homozygous in Pt 15 (heterozygosity rate)	BAC/PAC clone (chromosomal position)	Retention/ Deletion of clone on array	Array result confirmed by FISH
D8S1113 (60Mb)	Heterozygous (0.81)	RP11-114M5 (60Mb)	Retained	–
D8S1763 (60.3Mb)	Heterozygous (0.73)			
D8S1812 (60.8Mb)	Homozygous (0.70)			
		RP11-414L17 (61.7)	Deleted	Yes
		RP11-35A5 (61.8Mb)	Deleted	–
		RP3-491L6 (61.9Mb)	Deleted	–
		RP11-227F6 (62.4Mb)	Deleted	Yes
		RP11-45K10 (64.4Mb)	Deleted	–
		RP11-115G12 (65.4Mb)	Deleted	Yes
D8S512 (65.6Mb)	Homozygous (0.60)			
D8S544 (65.8Mb)	Homozygous (0.65)			
		RP1-252K6 (66Mb)	Retained	–

Figure 7.3.2.2 Location of microsatellite markers and BAC/PAC clones on chromosome 8q12 used for identification and validation of the extent of the deleted region.

7.4 Discussion

7.4.1 General discussion on the array-CGH results

This study demonstrates the increased sensitivity of array-CGH to detect deletions which may be missed by conventional cytogenetics using G-banding alone. This may be particularly true for deletions lying very close to the centromere where variations in the centromeric heterochromatin can make it difficult to resolve small changes in the adjacent G-bands.

The three patients identified with imbalances add weight to the use of array-CGH for the detection of constitutional mutations in patients with developmental malformations. Autosomal deletions, in almost all cases, cause a non-specific embryopathy that presents after birth as growth failure, mental retardation, and multiple malformations (Brewer et al. 1998). Based on the current literature, it appears that the detection sensitivity may be quite high for this technique if patients are strictly selected based on their phenotype and lack of positive genetic investigations (Vissers et al. 2003).

The identification of chromosome imbalances in 3/15 (20%) patients in this study is in keeping with the sensitivity detected by others in studies of children with dysmorphic features and learning difficulties (Shaw-Smith et al. 2004). One of the objectives for this study was to search for and identify candidate genes responsible for the 22q11DS phenotype. Some of the discussions below focus on this aim.

7.4.2 Discussion of results for specific patients

7.4.2.1 Identification of a 5Mb deletion on 5q11 in DNA from patient PS5

To the best of my knowledge this patient is the first patient with a deletion on 5q11.2 described in the literature (Prescott et al. 2005b).

It should be emphasised that although PS5 has many phenotypic features in common with 22q11DS, his facial features and overall phenotype are unique. However the

location of the deletion may identify genes involved in the development of the pharyngeal region and cardiac outflow tract that could associate with genes implicated in 22q11DS.

Chromosomal deletions often involve low copy repeats (LCRs) which are present in increased abundance around peri-centromeric regions. However, from the August 2001 build of the human genome there are no identified LCRs directly flanking the deleted region to account for the deletion in PS5 (Bailey et al. 2002).

Log_2 fluorescence ratios of -1 are theoretically expected for a deletion in patient DNA. In PS5 the values are in the -0.35 to -0.6 range (BAC numbers 3 and 5-8 in figure 7.3.1.1). There are several possible explanations for this. Firstly the deletion could be a polymorphism. This is unlikely as the clones did not show a reduced ratio in 20 control and 50 patient DNAs. The second possibility is that patient 5 is a mosaic for the deletion. Examination of ten metaphases for each absent FISH probe, and finding consistent heterozygosity throughout, made this hypothesis unlikely. Thirdly there is the possibility of segmental duplications on other chromosomes causing the increased ratios. This explanation is the most favoured as the deleted region is very close to the chromosome 5 centromere, a region that is known to harbour repetitive DNA sequences.

Table 7.4.2.1 shows a list of the 32 genes within the 6.5Mb potentially deleted region for this patient. There are several possible biologically relevant genes within the deleted region. For example, *Follistatin* is crucial in early embryonic development. *Follistatin* is expressed in the embryonic olfactory mesenchyme, cochlea epithelium, eye, tongue, stomach, intestine, pancreas, brain, spinal cord, blood vessels, heart and gonads (Feijen et al. 1994). *Follistatin* binds and inhibits *activin* which is a member of the transforming growth factor *TGFβ* superfamily of structurally related proteins. *Follistatin* binds *TGFβ* -like bone morphogenetic proteins (BMPs)-2,-4 and -7 and antagonizes BMP signalling by blocking their binding to the BMP receptors through forming a trimeric complex (Nakamura et al. 1990; Hemmati-Brivanlou et al. 1994; Iemura et al. 1998). This has parallels with *chordin* which also binds to and sequesters these BMPs, but with a higher affinity for BMP-2 and -4 (Piccolo et al. 1996). This is interesting in light of the fact that the *chordin* homozygous null mouse has many features of 22q11DS (Bachiller et al. 2003).

Gene	Protein	Mouse Model	Human mutations
NM_002202	ISLET-1 (insulin gene enhancer protein)	Null mice: embryonic lethal, mice have cardiac defects (Matzuk et al. 1995): see discussion.	Insulin dependent diabetes type 2 (Shimomura et al. 2000).
NM_015946	PELO; cell cycle regulation	Null mice: embryonic lethal, mice have defects in cell proliferation (Adham et al. 2003).	
NM_014100	Hypothetical protein		
NM_181501	INTEGRIN ALPHA 1 (Laminin and collagen receptor)	Null mice: develop normally, but develop osteoarthritis in later life (Zemmyo et al. 2003).	
NM_02203	INTEGRIN ALPHA 2 (Laminin and collagen receptor)	Null mice: defects in platelet adhesion to type 1 collagen and prolonged platelet aggregation (Chen et al. 2002).	
NM_004531	MOCS2; Molybdenum cofactor synthesis 2	Null mice: die within the 1st few days of life (Lee et al. 2002) .	Autosomal recessive: severe neurological defects and seizures (Reiss et al. 1999).
NM_006350	FST; Follistatin	Null mice: growth retarded, defects of the hard palate (Cai et al. 2003) :see discussion.	
NM_002495	NADH coenzyme Q reductase; mitochondrial respiratory chain complex 1		Autosomal recessive: severe neurological defects and seizures (van den et al. 1998).
NM_019087	Hypothetical protein		
Q9H385	Unknown		
NM_006308	HSPB3; Heat shock protein beta3		
NM_052870	SNAG1; Nexin associated glogi protein: involved in protein trafficking		
NM_007036	ESM1; Endothelial cell specific molecule		
NM_002104	GRANZYME K; released from cytotoxic cell lymphocytes and cause cell death		
NM_006144	GRANZYME A; released from cytotoxic cell lymphocyteS and cause cell death		
NM_152623	CDC20-like		
Q8TED1	Unknown. Protein homolgy to glutathione peroxidase; protection against oxidative stress		
NM_021147	UNG2; Uracil DNA glycosylase: involved with base excision repair	UNG null mice: develop B-cell lymphomas (Nilsen et al. 2003).	
NM_019030	DHX29; nucleic acid helicase		
NM_015360	K052_HUMAN; Helicase protein family		
NM_003711	PPAP2A; Phosphatidic acid phosphatase type 2A		
NM_173514	Protein with ambiguous function		
NM_024415	DEAD BOX PROTEIN 4; helicase involved in germ cell development (Tanaka et al. 2000) .		
NM_139017	GP130-LIKE MONOCYTE RECEPTOR		
NM_175767	IL6ST; IL6 beta chain precursor; GP130 signal transducer	Null mice: die before term with hypoplastic ventricular myocardium and reduced numbers of haematopoietic progenitors (Yoshida et al. 1996).	
NM_021104	RIBOSOMAL PROTEIN L41		
NM_024669	Protein; from ankyrin repeat protein family		
Q8N7E8	Unknown		
MAP3K1	MAP KINASE KINASE KINASE 1; involved with cell survival (Yujiri et al. 1998) .		
NM_153706	Protein with ambiguous function		
NM_152622	Protein with ambiguous function		
NM_022913	VASCULIN; possible role in atherogenesis (Bijnens et al. 2003).		

Table 7.4.2.1 List of genes in the deleted region in PS5 on 5q11.2.

A mouse model for *follistatin* has been developed with relevance to the phenotype of the deleted patient described in this paper. *Follistatin* null mice have hard palate defects and stunted growth (Matzuk et al. 1995). They also have decreased respiratory muscle mass, shiny taut skin, defects in the thirteenth pair of ribs and a decrease in the number of lumbar vertebrae. In addition their whisker and tooth development is abnormal and they die within a few hours of birth from respiratory distress. As patient PS5 has a bifid uvula, velopharyngeal insufficiency, abnormal dentition and short stature, it is possible these are due to haploinsufficiency of the gene for *FOLLISTATIN*.

A second gene within the deleted region is *ISLET-1* which encodes a member of the LIM/homeodomain family of transcription factors that binds to the enhancer region of the insulin gene. Interestingly mice that are homozygous null for *islet-1* exhibit growth retardation and have no outflow tract, no right ventricle and absence of much of the atria (Cai et al. 2003). Recent work has demonstrated that *Islet-1* is upstream of *Shh* in a pathway required for cardiac morphogenesis (Lin et al. 2006). Haploinsufficiency of this gene in patient PS5 could contribute to his heart defect of tetralogy of Fallot in which the pulmonary outflow tract and right ventricle are abnormal. However mutation analysis of *ISLET-1* in non syndrome tetralogy of Fallot patients (unpublished data from MMU) showed no mutations. *Islet-1* has some downstream target genes in common with *follistatin* (BMPs-2,-4 and -7) and *Tbx1* (fibroblast growth factors 8 and 10) (Vitelli et al. 2002b), and hypomorphic alleles of *Fgf-8* interact epistatically with *Tbx1* (Vitelli et al. 2002b).

Although *islet-1* and *follistatin* heterozygous mice are phenotypically normal on the backgrounds reported, there are many examples of differences in gene dosage effects between man and mouse. Additionally, hemizygosity of both genes may be required for the phenotype, via an epistatic interaction. Such an interaction could explain the main features of the phenotype in patient PS5 - namely the heart and palate defects and the short stature. However *Islet1/+;Tbx1/+* transheterozygous murine crosses showed no change in phenotype (Antonio Baldini personal communication).

Many of the other genes within the deleted region may also have roles in development such as the helicases and *MAP KINASE KINASE KINASE 1*, *PELO* and *ECM-1*. These

genes have a more general function in development so if they contribute to the phenotype in patient PS5, their role is less clear (Table 7.4.2.1).

7.4.2.2 Identification of a 5-5.7Mb deletion on 8q12 in DNA from patient PS15

To the best of my knowledge there are no LCRs directly flanking the deleted region in patient PS15 (Bailey et al. 2002). However work is ongoing to look in more detail at this region by Dr B. de Vries, Nijmegen (personal communication).

The phenotype seen in PS15 is more severe than that seen in most patients with 22q11DS as the patient had an absent corpus callosum and no speech. The heart phenotype of a type A aortic arch interruption and coarctation is unlikely to represent a defect in 4th pharyngeal arch artery development as these defects usually result in type B interruption. The phenotype is therefore quite distinct from 22q11DS.

Since the deletion in this patient was ascertained, Vissers *et al* published work identifying the chromodomain gene *CHD7* as the causative gene in CHARGE syndrome (Vissers et al. 2004). This gene was identified by establishing a shortest region of deletion overlap by studying the DNA from CHARGE patients with array-CGH. Two patients with deletions and one patient with a translocation narrowed down the region to 2.3Mb allowing screening of nine candidate genes. Ten heterozygous mutations were found in *CHD7* including seven stop mutations, two missense mutations and one mutation at an exon-intron boundary. The patients with the deletion in the 8q12 region both had coloboma, heart malformations, growth retardation and development and ear abnormalities including deafness (Vissers et al. 2004).

Although PS15 had some features in common with CHARGE syndrome she did not have coloboma or choanal atresia. Thus *CHD7* haploinsufficiency may be associated with phenotypes other than CHARGE syndrome.

Inspection of the other 12 genes within the region did not reveal any obvious candidate genes to account for the phenotype. *SDCBP* is a gene 1Mb proximal to the deletion and is known to bind to *SOX4* (Geijsen et al. 2001). *Sox4* deficient mice have been shown to have aortic arch and OFT defects (Schilham et al. 1996). Although *SDCBP* is not

directly disrupted by the deletion, it is possible that its expression could be disrupted by a position effect.

7.4.2.3 Identification of a chromosomal imbalance on 5p/5q in DNA from patient PS25

It can be seen from figure 7.3.1.3 that the hybridisation pattern of PS25 is quite "noisy" with many clones giving a log2 ratio beyond 0. This could represent DNA with impurities. Note that not all of the BACs within the proposed 5p deleted region are present at a $\log_2$ ratio between -0.5 and -1. At least one of these BACs RP11-13H10 is a known "problematic clone" indicating it is from a region of a copy number polymorphism between the test and reference genomes (figure 7.3.1.3 green circle). Validation experiments are needed for this patient to confirm the extent of the deletion.

One would expect clones that are duplicated to have a $\log_2$ hybridisation ratio in the order of +0.58 ($\log_2$ 3/2=0.58). Several of the clones in PS25 at the 5q telomere are in this range (BACs 9-12 figure 7.3.1.3). There are a few BACs after this cluster that don't appear to be significantly changed. Again this could be due to "problematic clones" but it could also indicate that the duplication on 5q is interstitial. The former argument is more likely given that there is a 5p deleted region as well which could be the result of a pericentric inversion in a parent (figure 7.4.2.3.1). The resultant imbalance in PS25 suggested by the array-CGH result is trisomy 5q35.1-qter and monosomy 5p15.31-pter (figure 7.4.2.3.1).

Deletions on chromosome 5p lead to a variety of developmental defects with most cases classified as cri du chat syndrome (OMIM 123450; Niebuhr 1978). Cri du chat syndrome has several phenotypic components, including the characteristic cat like cry that gives the syndrome its name, characteristic facial features such as hypertelorism, severe speech delay and learning difficulties. The deletions can be terminal or interstitial and occasionally occur in the context of a cytogenetically complex karyotype (Sreekantaiah et al. 1999).

Recently Zhang *et al* performed a detailed study of phenotype-genotype relationships by analysing DNA copy number changed in 94 patients with cri du chat by using array-CGH (Zhang et al. 2005b). The group refined the "cry" region for cri du chat to a 6-

7.5Mb region on 5p15.31. It can be seen from figure 7.4.2.3.3 that the proximal boundary for the "cry" region lies within the deletion breakpoint for PS25. Knowledge of the "cry" phenotype in PS25 and finer mapping of the breakpoint could have refined the "cry" region for cri du chat. Unfortunately it was impossible to ascertain this information for this project.

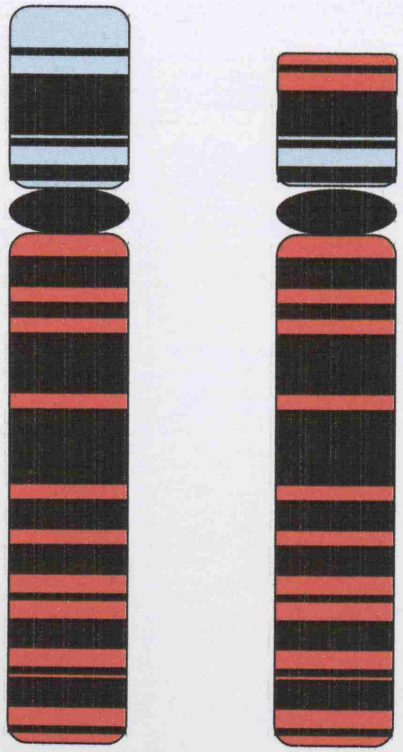


Figure 7.4.2.3.1 .
Ideogram of chromosome 5s to represent the possible chromosomal imbalance of patient PS25 with trisomy 5q35.1-qter and monosomy 5p15.3-pter. Pale blue: p arm; red: q arm.

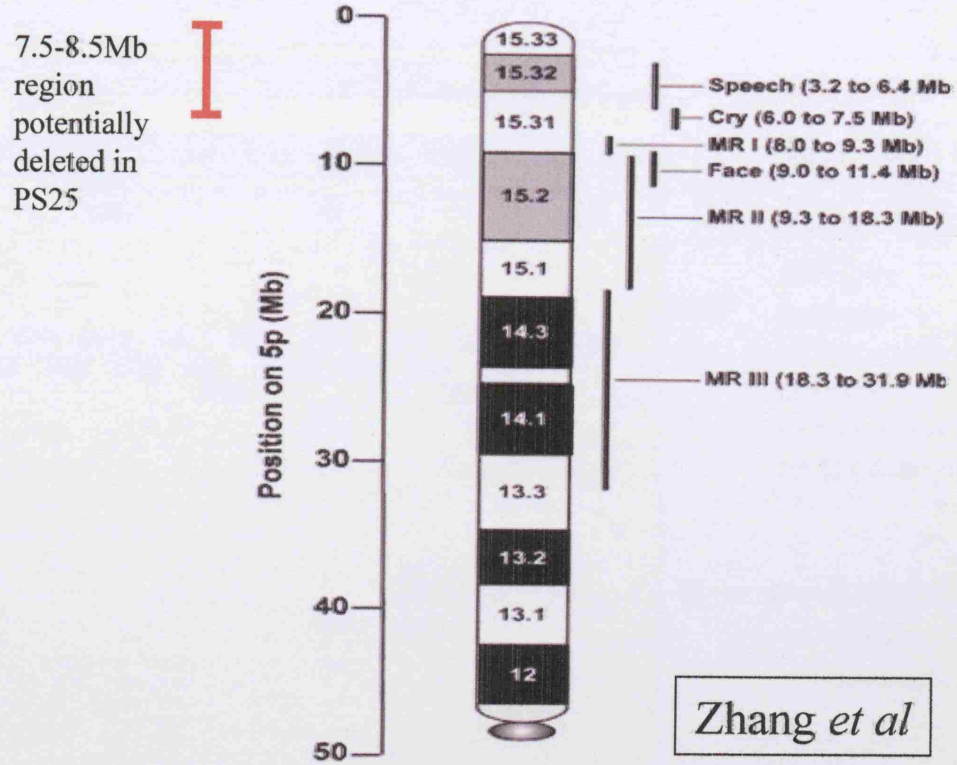


Figure 7.4.2.3.2
Summary of phenotype-genotype relationships in 94 Cri du Chat patients from Zhang *et al*. Red line depicts the extent of the deleted region in PS25 from the array-CGH results. Patient PS25 is deleted for the "speech" region and for part of the "cry" region. Also the deletion slightly overlaps with the "MR1" region which is "mental retardation 1".Adapted from Zhang *et al.*

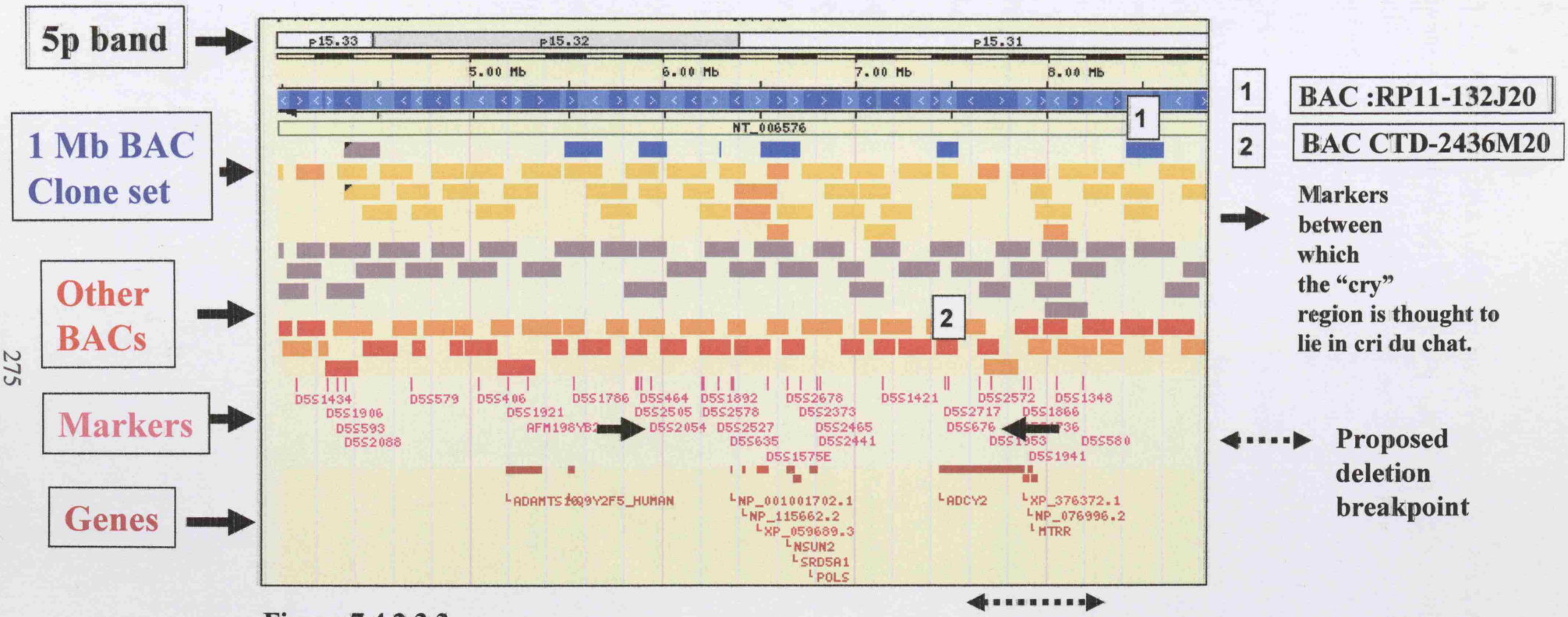


Figure 7.4.2.3.3

Web page from Ensemble v34-October 2005 (http://www.ensembl.org/Homo_sapiens/contigview) to show the region on 5p15.31 thought to be deleted in PS25 . BAC clone RP11-132J20 was not deleted on the array and clone CTD-2436M20 had a reduced ratio on the array consistent with a deletion. The microsatellite markers signified with black arrows are the markers between which the "cry" region in cri du chat is thought to lie (Zhang *et al*). Knowledge of the "cry" phenotype in PS25 and finer mapping of the breakpoint could have refined the "cry" region for cri du chat.

Also lacking is information about the degree of development of PS25. When the patient was referred he was only a baby, so this data was not available. The estimated 7.5-8.5 Mb deletion on 5p would just encroach on the "MR1" or mental retardation 1 region proposed by Zhang *et al.*

Patients have also been described in the literature with 5q35-qter duplications. Such patients have a characteristic phenotype of craniosynostosis, mental retardation, specific facial dysmorphism, and mild acral skeletal features. The Hunter-McAlpine syndrome results from such duplications (Hunter et al. 2005).

7.4.3 The future of Array-CGH and its role in diagnostic and research medicine

Array-CGH screening of a selective patient group increases capacity to detect submicroscopic chromosome number changes and therefore the ability to identify genes responsible for developmental disorders. Once the arrays have been developed and verified, array-CGH also has the advantage of being a relatively quick procedure with high throughput. On identification of an aberration it is fairly straightforward to relate this copy number change to the genomic sequence. The main drawback will be the identification of many imbalances where the pathogenicity is unproven. Despite this, array-CGH has many diagnostic, prognostic, counselling and research applications and will undoubtedly be lucrative in identifying pathways of disease.

Chapter 8 Discussion and future work

8.1 Yeast two hybrid analysis of Tbx1 (chapter 3)

The aim of this experiment was to identify potential protein binding partners for Tbx1. As Tbx1 has only recently been identified as the main gene involved in the 22q11DS phenotype, little work has been done to identify potential protein interactors. It would be anticipated that as a transcriptional activator it is likely to have protein binding partners that act as cofactors in transcription. It is known that other T-box proteins have protein binding partners that are important for their transcriptional activity. There are currently just two T-box genes with published yeast two hybrid data: Tbx2 and Tbx5. Yeast two hybrid analysis of TBX-2 has shown its interaction with components of the *c.elegans* SUMO-conjugation pathway (Chowdhuri et al. 2006). Krasue *et al* used the C terminus of TBX5 as bait and identified a PDZ-LIM protein important in limb and heart development named LMP-4 (Krause et al. 2004). This was also found to interact with TBX4 (Krause et al. 2004).

The full length Tbx1 protein caused autoactivation of the reporter gene(s), prohibiting its use in a two hybrid screen. The C terminal construct produced a high yield of positive clones, but only a small proportion of these had identifiable plasmid prey cDNA. The sequencing results revealed only likely false positive interactors.

The N terminal construct had more promising results with four independent clones of the cytoplasmic dynein light chain 1 (DLC1) identified that were in frame with the VP16 activation domain. On first inspection this candidate would appear to be a false positive as it is mainly cytoplasmic. However, further analysis of the literature revealed its interaction with several transcription factors and a possible role in chaperoning proteins between the cytoplasm and the nucleus. As Tbx1 has also been observed in the cytoplasm as well as the nucleus, it is possible that DLC1 serves to sequest Tbx1 in the cytoplasm and assist in its transport to the nucleus when upstream signals are activated.

8.1.1 Future work

There are many avenues of follow-up work that need to be performed to verify DLC1 as a true interactor with Tbx1. Subcloning DLC1 into the pGBDU vector and subcloning N terminal mTbx1 into the VP16 vector, followed by retransformation into yeast to try

to recapitulate the reporter gene phenotype is one possible experiment to test the interaction. Different truncations of both proteins could be manipulated to identify the specific domains responsible for the interaction.

GST (glutathione-S-transferase) capture assay experiments provide a direct *in vitro* method of analysing the interaction and would be the next step in examining the candidate protein. This method is one of the most widely used techniques with which to investigate protein-protein interactions. It is an affinity chromatography method that involves using a tagged or labelled bait to create a specific affinity matrix that will enable binding and purification of a prey protein from a lysate sample. Typically the bait protein is a recombinant GST fusion protein produced in *E. coli*, and the prey protein is ^{35}S-labelled by *in vitro* translation. The GST-fusion protein bound to glutathione agarose beads is incubated with the radio-labelled prey protein in the presence of a competitor protein to block non-specific interactions. Bound proteins are then washed and the capture of prey protein is determined by SDS-PAGE and autoradiography

Co-immunoprecipitation studies would provide further evidence for an *in vivo* interaction. This method is designed to affinity purify a bait protein antigen together with its binding partner using a specific antibody against the bait. The protein/antigen complex is immobilised on a solid support such antibody binding beads and then eluted and analysed.

Immunohistochemistry and flourescence resonance energy transfer studies can also be employed to examine the cellular localisation of the proteins to see if they co-localise and under which specific conditions.

8.2 Identification and verification of Df1 target genes (chapters 4 and 5)

22q11 deletion syndrome results in a phenotype that affects a specific area of development: the pharyngeal apparatus. The development of the *Df1* mouse model has provided a mammalian system within which it is possible to study the gene expression changes that are occurring at the time of pharyngeal arch and pouch development.

The availability of the Affymetrix microarray facility in house provided the opportunity of utilising a modern technology to examine expression changes at a specific time period. Use of a heterozygous mouse model was anticipated to be at the limits of gene expression change detection, as estimated gene changes were likely to be less than two fold. However use of the *Df1* mouse model has a distinct advantage over most other mouse mutants in that there are 22 genes within the deleted region that can be exploited to assess the reliability of expression data. These genes were used as "in built" positive controls and allowed quite stringent criteria when performing filtering and statistical analysis of data.

The primary aim of the experiment was to look for gene expression changes within the branchial arch region that might be responsible for the mouse phenotype (i.e. direct or indirect targets of the deletion). Studying a deleted mouse model also provided other avenues for research, besides "target" identification. It provided the opportunity to look for the effect of expression changes of genes adjacent to, but not within the deletion. This is often an important question in clinical and research discussions. Secondly the model allowed analysis of whether hemizygosity for *Df1* genes resulted in a significantly greater than two-fold change in the expression level of any other gene represented on the array. Thirdly relative expression analysis was used to see if there was any dosage compensation by the non-deleted allele. Dosage compensation could partly explain the lack of haploinsufficient phenotypes for the majority of genes within the genome.

Verification was provided by RTQPCR and whole mount *in situ* hybridisation experiments. Similar to microarray technology, detection of expression changes less than two fold is pushing the limits of RTQPCR. Generally the microarray RTQPCR results for the *Df1* mouse have provided some potential candidates for further analysis, but the precise magnitude of change is difficult to quantify. Such subtle expression changes would also be unlikely to show up on whole mount *in situ* hybridisation analysis.

Utilisation of the Affymetrix microarray technology allowed the aims to be achieved with 9 out of 12 *Df1* deleted genes present on the microarrays passing the stringent statistical filtering applied. Several genes involved in vasculogenesis and cardiogenesis were validated by RTQPCR including *Connexin 45*, a gene required for normal vascular

development and *Dnajb9*, a gene implicated in microvascular differentiation. In addition a known *Tbx1* target *FoxA2*, passed statistical filtering. There was no evidence of any dosage compensation of deleted genes, suggesting this phenomenon is rare, and no dysregulation of genes mapping immediately adjacent to the deletion was detected. *Crkl*, another gene implicated in the 22q11DS phenotype (Guris et al. 2006) was found to be downregulated by microarray and RTQPCR.

8.2.1 Future Work

This study compared the expression profile of an E10.5 mouse embryo deleted for 22 genes with a wild type mouse embryo. Comments can therefore be made on general gene changes, but it in such a system it is impossible to identify which deleted genes are responsible for the specific expression changes. A much purer experiment would be to use the *Tbx1* mouse model for microarray analysis. This has already been done (Ivins et al. 2005) and the problems faced included the fact that there weren't many significantly changes genes when *Tbx1* heterozygous mice were used (Dr Ivins: personal communication). When the homozygous model was used there was tissue loss due to hypoplasia of the pharyngeal apparatus resulting in the potential for false positive targets (Ivins et al. 2005).

Further work is ongoing in the MMU laboratory to overcome these problems by specifically isolating *Tbx1* expressing cells from embryos through fluorescence activated cell sorting analysis and then extracting RNA for microarray analysis. Potential targets can then be examined in *Tbx1* inducible cell lines that might magnify the effect of the change and where experimental conditions are more homogeneous.

Follow up would ultimately involve animal transheterozygote crosses between *Df1* mice and mice with the candidate gene either under or over expressed. An examination of the phenotype to see whether there is rescue or exacerbation would answer the question of whether they were involved in the same developmental pathway. Direct target identification can be investigated by utilising the techniques of chromatin immunoprecipitation and luciferase reporter assays.

8.3 Identification and verification of *Pax3* target genes

Initially the *Pax3* microarray screen was commenced as an introduction to the technique as the *Df1* mouse colony was delayed in its arrival due to the Baylor flood (Schub 2002). However, generally the results were more promising than for the *Df1* experiment, although time constraints meant that follow up work was minimal. This is perhaps due to the use of a homozygous (Sp^{2H}/Sp^{2H}) rather than a heterozygous (*Df/+*) model exhibiting more severe defects and therefore larger gene expression changes. Clearly along with this advantage is the problem of tissue loss and potential false positive results.

Amongst the results was a known *Pax3* target (*lunatic fringe*) as well as expected changed genes such as *Hira* and *Hmx1* which are expressed in neural crest cells (*Pax3* deficient mice have reduced neural crest cell numbers). Along with these genes came many plausible *Pax3* targets including *Wnt5a* which was verified by RTQPCR and whole mount *in situ* hybridisation..

8.3.1 Future Work

Clearly more of the potential targets need to be tested by RTQPCR analysis. Whole mount *in situ* hybridisation work together with sectioning would allow a more thorough inspection of potentially changed genes, especially if the change was suspected to be due to tissue loss. Analysis of upstream segments of potential targets for *Pax3* binding sites would add credence to candidate genes.

8.4 Investigation of patients by array-CGH referred as suspected cases of 22q11DS

A collaboration with Nigel Carter at the Sanger Centre allowed the rapidly expanding new technique of array-CGH to be applied to some of the patient samples that were not deleted for the *Tuple* FISH probe. This yielded three patients with chromosomal deletions that I was able to further analyse to verify the array-CGH result. One patient had a novel deletion at 5q11 which encompassed several potential candidate genes for the tetralogy of Fallot heart phenotype in this patient. A second patient carried a deletion involving the *CHD7* gene which has recently been found to be mutated in patients with CHARGE syndrome. It is likely that with the wide expansion of array-

CGH that many more genes involved in developmental biology will be identified through shortest region of overlap studies.

8.4.1 Future Work

As exemplified by Vissers *et al*, gene identification by array-CGH is an important goal (Vissers et al. 2004). If other patients with deletions at 5q11 are identified this might allow refinement of a region of overlap and minimisation of candidate genes for the phenotype. Mutation analysis of candidate genes (particularly ISL1) in patients with part of the phenotype such as tetralogy of Fallot, might identify disease causing genes.

As it is now thought that small submicroscopic deletions and duplications constitute up to 15% of all mutations underlying monogenic diseases (Vissers et al. 2005). The resolving power of array-CGH is currently improving from the megabase to the kilobase level (Ishkanian et al. 2004). Application of non-deleted patients with a 22q11 deletion phenotype to this higher resolution array would allow a more refined analysis of these patients for submicroscopic genome imbalances.

Other avenues of exploration of non-deleted patients with an overlapping phenotype with 22q11DS, would be SNP-chip analysis of patients from consanguineous families to look for loci for homozygosity by descent. In the future application of patient DNA to microarray sequencing chips containing genes involved in heart development might identify the causative gene in some patients.

References

1. Abu-Elmagd,M., Ishii,Y., Cheung,M., Rex,M., Le Rouedec,D. & Scotting,P.J. 2001. cSox3 expression and neurogenesis in the epibranchial placodes. *Dev. Biol.* 237, 258-269.

2. Abu-Issa,R., Smyth,G., Smoak,I., Yamamura,K. & Meyers,E.N. 2002. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. *Development* 129, 4613-4625.

3. Adachi,M., Tachibana,K., Masuno,M., Makita,Y., Maesaka,H., Okada,T., Hizukuri,K., Imaizumi,K., Kuroki,Y., Kurahashi,H. & Suwa,S. 1998. Clinical characteristics of children with hypoparathyroidism due to 22q11.2 microdeletion. *European Journal of Pediatrics* 157, 34-38.

4. Adham,I.M., Sallam,M.A., Steding,G., Korabiowska,M., Brinck,U., Hoyer-Fender,S., Oh,C. & Engel,W. 2003. Disruption of the pelota gene causes early embryonic lethality and defects in cell cycle progression. *Mol. Cell Biol.* 23, 1470-1476.

5. Afouda,B.A., Ciau-Uitz,A. & Patient,R. 2005. GATA4, 5 and 6 mediate TGFbeta maintenance of endodermal gene expression in Xenopus embryos. *Development* 132, 763-774.

6. Agulnik,S.I., Papaioannou,V.E. & Silver,L.M. 1998. Cloning, mapping, and expression analysis of TBX15, a new member of the T-Box gene family. *Genomics* 51, 68-75.

7. Agulnik,S.I., Ruvinsky,I. & Silver,L.M. 1997. Three novel T-box genes in Caenorhabditis elegans. *Genome* 40, 458-464.

8. Altschul,S.F., Gish,W., Miller,W., Myer,E.W. & Lipman,D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.

9. Alwine,J.C., Kemp,D.J. & Stark,G.R. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. U. S. A* 74, 5350-5354.

10. Amati,F., Conti,E., Novelli,A., Bengala,M., Diglio,M.C., Marino,B., Giannotti,A., Gabrielli,O., Novelli,G. & Dallapiccola,B. 1999. Atypical deletions suggest five 22q11.2 critical regions related to the DiGeorge/velo-cardio-facial syndrome. *Eur J Hum Genet* 7, 903-909.

11. Andachi,Y. 2004. Caenorhabditis elegans T-box genes tbx-9 and tbx-8 are required for formation of hypodermis and body-wall muscle in embryogenesis. *Genes Cells* 9, 331-344.

12. Arnold,J.S., Werling,U., Braunstein,E.M., Liao,J., Nowotschin,S., Edelmann,W., Hebert,J.M. & Morrow,B.E. 2006. Inactivation of Tbx1 in the pharyngeal endoderm results in 22q11DS malformations. *Development* 133, 977-987.

13. Ataliotis,P., Ivins,S., Mohun,T.J. & Scambler,P.J. 2005. XTbx1 is a transcriptional activator involved in head and pharyngeal arch development in Xenopus laevis. *Developmental Dynamics* 232, 979-991.

14. Auerbach,R. Analysis of the developmental effects of a lethal mutation in the house mouse. J.Exp.Zool 127, 305-329. 1954.

15. Ayuk,P.T., Sibley,C.P., Donnai,P., D'Souza,S. & Glazier,J.D. 2000. Development and polarization of cationic amino acid transporters and regulators in the human placenta. *Am. J. Physiol Cell Physiol* 278, C1162-C1171.

16. Bachiller,D., Klingensmith,J., Shneyder,N., Tran,U., Anderson,R., Rossant,J. & De Robertis,E.M. 2003. The role of chordin/Bmp signals in mammalian pharyngeal development and DiGeorge syndrome. *Development* 130, 3567-3578.

17. Baglin,T.P., Carrell,R.W., Church,F.C., Esmon,C.T. & Huntington,J.A. 2002. Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism. *Proc. Natl. Acad. Sci. U. S. A* 99, 11079-11084.

18. Bailey,J.A., Gu,Z., Clark,R.A., Reinert,K., Samonte,R.V., Schwartz,S., Adams,M.D., Myers,E.W., Li,P.W. & Eichler,E.E. 2002. Recent segmental duplications in the human genome. *Science* 297, 1003-1007.

19. Baldwin,C.T., Hoth,C.F., Amos,J.A., da Silva,E.O. & Milunsky,A. 1992. An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* 355, 637-638.

20. Bamshad,M., Lin,R.C., Law,D.J., Watkins,W.C., Krakowiak,P.A., Moore,M.E., Franceschini,P., Lala,R., Holmes,L.B., Gebuhr,T.C., Bruneau,B.G., Schinzel,A., Seidman,J.G., Seidman,C.E. & Jorde,L.B. 1997. Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nat. Genet.* 16, 311-315.

21. Basson,C.T., Bachinsky,D.R., Lin,R.C., Levi,T., Elkins,J.A., Soults,J., Grayzel,D., Kroumpouzou,E., Traill,T.A., Leblanc-Straceski,J., Renault,B., Kucherlapati,R., Seidman,J.G. & Seidman,C.E. 1997. Mutations in human TBX5 cause limb and cardiac malformation in Holt-oram syndrome. *Nature Genetics* 15, 30-35.

22. Bayes,M., Magano,L.F., Rivera,N., Flores,R. & Perez Jurado,L.A. 2003. Mutational mechanisms of Williams-Beuren syndrome deletions. *Am. J. Hum. Genet.* 73, 131-151.

23. Bearden,C.E., Woodin,M.F., Wang,P.P., Moss,E., McDonald-McGinn,D., Zackai,E., Emannuel,B. & Cannon,T.D. 2001. The neurocognitive phenotype of the 22q11.2 deletion syndrome: selective deficit in visual-spatial memory. *J. Clin. Exp. Neuropsychol.* 23, 447-464.

24. Beechey,C. Mutations at the Sp locus. Searle, A. G. Mouse News Lett 75, 28. 1986.

25. Bernasconi,M., Remppis,A., Fredericks,W.J., Rauscher,F.J., III & Schafer,B.W. 1996. Induction of apoptosis in rhabdomyosarcoma cells through

down-regulation of PAX proteins. *Proc. Natl. Acad. Sci. U. S. A* 93, 13164-13169.

26. Berti,L., Mittler,G., Przemeck,G.K., Stelzer,G., Gunzler,B., Amati,F., Conti,E., Dallapiccola,B., Hrabe,d.A., Novelli,G. & Meisterernst,M. 2001. Isolation and characterization of a novel gene from the DiGeorge chromosomal region that encodes for a mediator subunit. *Genomics* 74, 320-332.

27. Bertolotto,C., Abbe,P., Hemesath,T.J., Bille,K., Fisher,D.E., Ortonne,J.P. & Ballotti,R. 1998. Microphthalmia gene product as a signal transducer in cAMP-induced differentiation of melanocytes. *JCB* 142, 827-835.

28. Bobanovic,L.K., Royle,S.J. & Murrell-Lagnado,R.D. 2002. P2X receptor trafficking in neurons is subunit specific. *J. Neurosci.* 22, 4814-4824.

29. Bondurand,N., Pingault,V., Goerich,D.E., Lemort,N., Sock,E., Caignec,C.L., Wegner,M. & Goossens,M. 2000. Interaction among SOX10, PAX3 and MITF, three genes altered in Waardenburg syndrome. *Human Molecular Genetics* 9, 1907-1917.

30. Borycki,A.G., Li,J., Jin,F., Emerson,C.P. & Epstein,J.A. 1999. Pax3 functions in cell survival and in pax7 regulation. *Development* 126, 1665-1674.

31. Botta,A., Lindsay,E.A., Jurecic,V. & Baldini,A. 1997. Comparative mapping of the DiGeorge syndrome region in mouse shows inconsistent gene order and diferential degree of gene conservation. *Mammalian Genome* 8, 890-895.

32. Boulter,C., Mulroy,S., Webb,S., Fleming,S., Brindle,K. & Sandford,R. 2001. Cardiovascular, skeletal, and renal defects in mice with a targeted disruption of the Pkd1 gene. *Proc. Natl. Acad. Sci. U. S. A* 98, 12174-12179.

33. Braybrook,C., Doudney,K., Marcano,A.C., Arnason,A., Bjornsson,A., Patton,M.A., Goodfellow,P.J., Moore,G.E. & Stanier,P. 2001. The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. *Nat. Genet.* 29, 179-183.

34. Bressan,A., Somma,M.P., Lewis,J., Santolamazza,C., Copeland,N.G., Gilbert,D.J., Jenkins,N.A. & Lavia,P. 1991. Characterization of the opposite-strand genes from the mouse bidirectionally transcribed HTF9 locus. *Gene* 103, 201-209.

35. Brewer,C., Holloway,S., Zawalnyski,P., Schinzel,A. & FitzPatrick,D. 1998. A chromosomal deletion map of human malformations. *Am. J. Hum. Genet.* 63, 1153-1159.

36. Brown,C.B., Wenning,J.M., Lu,M.M., Epstein,D.J., Meyers,E.N. & Epstein,J.A. 2004. Cre-mediated excision of Fgf8 in the Tbx1 expression domain reveals a critical role for Fgf8 in cardiovascular development in the mouse. *Dev. Biol.* 267, 190-202.

37. Bruneau,B.G., Nemer,G., Schmitt,J.P., Charron,F., Robitaille,L., Caron,S., Conner,D.A., Gessler,M., Nemer,M., Seidman,C.E. & Seidman,J.G. 2001. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* 106, 709-721.

38. Buccoliero,R., Bodennec,J. & Futerman,A.H. 2002. The role of sphingolipids in neuronal development: lessons from models of sphingolipid storage diseases. *Neurochemical Res* 27, 565-574.

39. Buckley,P.G., Mantripragada,K.K., Benetkiewicz,M., Tapia-Paez,I., Diaz,D.S., Rosenquist,M., Ali,H., Jarbo,C., De Bustos,C., Hirvela,C., Sinder,W.B., Fransson,I., Thyr,C., Johnsson,B.I., Bruder,C.E., Menzel,U., Hergersberg,M., Mandahl,N., Blennow,E., Wedell,A., Beare,D.M., Collins,J.E., Dunham,I., Albertson,D., Pinkel,D., Bastian,B.C., Faruqi,A.F., Lasken,R.S., Ichimura,K., Collins,V.P. & Dumanski,J.P. 2002. A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. *Human Molecular Genetics* 11, 3221-3229.

40. Budarf,M.L., Collins,J., Gong,W., Roe,B., Wang,Z., Bailey,L.C., Sellinger,B., Michaud,D., Driscoll,D.A. & Emanuel,B.S. 1995. Cloning a balanced translocation associated with DiGeorge syndrome and identification of a disrupted candidate gene. *Nature Genetics* 10, 269-278.

41. Bush,J.O., Lan,Y. & Jiang,R. 2004. The cleft lip and palate defects in Dancer mutant mice result from gain of function of the Tbx10 gene. *Proc. Natl. Acad. Sci. U. S. A* 101, 7022-7027.

42. Cai,C.L., Liang,X., Shi,Y., Chu,P.H., Pfaff,S.L., Chen,J. & Evans,S. 2003. Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev. Cell* 5, 877-889.

43. Cai,D.H. & Brauer,P.R. 2002. Synthetic matrix metalloproteinase inhibitor decreases early cardiac neural crest migration in chicken embryos. *Developmental Dynamics* 224, 441-449.

44. Candille,S.I., van Raamsdonk,C.D., Chen,C., Kuijper,S., Chen-Tsai,Y., Russ,A., Meijlink,F. & Barsh,G.S. 2004. Dorsoventral patterning of the mouse coat by Tbx15. *PLoS. Biol.* 2, E3.

45. Carey,A.H., Kelly,D., Halford,S., Wadey,R., Wilson,D., Goodship,J., Burn,J., Paul,T., Sharkey,A., Dumanski,J., Nordenskjold,M., Williamson,R. & Scambler,P.J. 1992. Molecular genetic study of the frequency of monosomy 22q11 in DiGeorge syndrome. *Am. J. Hum. Genet.* 51, 964-970.

46. Carlson,C., Sirotkin,H., Pandita,R., Goldberg,R., McKie,J., Wadey,R., Patanjali,S.R., Weissman,S.M., Anyane-Yeboa,K., Warburton,D., Scambler,P., Shprintzen,R., Kucherlapati,R. & Morrow,B.E. 1997. Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am. J. Hum. Genet.* 61, 620-629.

47. Carreira,S., Dexter,T.J., Yavuzer,U., Easty,D.J. & Goding,C.R. 1998. Brachyury-related transcription factor Tbx2 and repression of the melanocyte-specific TRP-1 promoter. *Mol Cell Biol* 18, 5099-5108.

48. Carrozzo,R., Rossi,E., Christian,S.L., Kittikamron,K., Livieri,C., Corrias,A., Pucci,L., Fois,A., Simi,P., Bosio,L., Beccaria,L., Zuffardi,O. & Ledbetter,D.H. 1997. Inter- and intrachromosomal rearrangements are both involved in the

origin of 15q11-q13 deletions in Prader-Willi syndrome. *Am. J. Hum. Genet.* 61, 228-231.

49. Carson,C.T., Kinzler,E.R. & Parr,B.A. 2000. Tbx12, a novel T-box gene, is expressed during early stages of heart and retinal development [In Process Citation]. *Mechanisms of Development* 96, 137-140.

50. Carter,M.G., Johns,M.A., Zeng,X., Zhou,L., Zink,M.C., Mankowski,J.L., Donovan,D.M. & Baylin,S.B. 2000. Mice deficient in the candidate tumor suppressor gene Hic1 exhibit developmental defects of structures affected in the Miller-Dieker syndrome. *Human Molecular Genetics* 9, 413-419.

51. Chalepakis,G., Goulding,M., Read,A., Strachan,T. & Gruss,P. 1994a. Molecular basis of splotch and Waardenburg Pax-3 mutations. *Proc. Natl. Acad. Sci. U. S. A* 91, 3685-3689.

52. Chalepakis,G., Jones,F.S., Edelman,G.M. & Gruss,P. 1994b. Pax3 contains domains for transcription activation and transcription inhibition. *Proc. Natl. Acd. sci USA* 91, 12745-12749.

53. Chalepakis,G., Stoykova,A., Wijnholds,J., Tremblay,P. & Gruss,P. 1993. Pax: gene regulators in the developing nervous system. *J. Neurobiol.* 24, 1367-1384.

54. Chan,W.Y., Cheung,C.S., Yung,K.M. & Copp,A.J. 2004. Cardiac neural crest of the mouse embryo: axial level of origin, migratory pathway and cell autonomy of the splotch (Sp2H) mutant effect. *Development* 131, 3367-3379.

55. Chapman,D.L., Garvey,N., Hancock,S., Alexiou,M., Agulnik,S.I., Gibson-Brown,J.J., Cebra-Thomas,J., Bollag,R.J., Silver,L.M. & Papaioannou,V.E. 1996. Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Developmental Dynamics* 206, 379-390.

56. Chapman,D.L. & Papaioannou,V.E. 1998. Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature* 391, 695-697.

57. Chelsey,P. 1935. Development of the *short-tailed* mutant in the house mouse. *J. Exp. Zool* 70, 429-459.

58. Chen,E., Hermanson,S. & Ekker,S.C. 2004. Syndecan-2 is essential for angiogenic sprouting during zebrafish development. *Blood* 103, 1710-1719.

59. Chen,J., Diacovo,T.G., Grenache,D.G., Santoro,S.A. & Zutter,M.M. 2002. The alpha(2) integrin subunit-deficient mouse: a multifaceted phenotype including defects of branching morphogenesis and hemostasis. *Am. J. Pathol.* 161, 337-344.

60. Chen,J., Zhong,Q., Wang,J., Cameron,R.S., Borke,J.L., Isales,C.M. & Bollag,R.J. 2001. Microarray analysis of Tbx2-directed gene expression: a possible role in osteogenesis. *Mol. Cell Endocrinol.* 177, 43-54.

61. Chen,W.Y., Zeng,X., Carter,M.G., Morrell,C.N., Chiu Yen,R.W., Esteller,M., Watkins,D.N., Herman,J.G., Mankowski,J.L. & Baylin,S.B. 2003. Heterozygous disruption of Hic1 predisposes mice to a gender-dependent spectrum of malignant tumors. *Nat. Genet.* 33, 197-202.

62. Chiang,C., Litingtung,Y., Lee,E., Young,K.E., Coden,J.L., Westphal,H. & Beachy,P.A. 1996. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* 383, 407-413.

63. Chiang,P.W. & Kurnit,D.M. 2003. Study of dosage compensation in Drosophila. *Genetics* 165, 1167-1181.

64. Chieffo,C., Garvey,N., Gong,W., Roe,B., Zhang,G., Silver,L., Emanuel,B.S. & Budarf,M.L. 1997. Isolation and characterization of a gene from the DiGeorge chromosomal region homologous to the mouse Tbx1 gene. *Genomics* 43, 267-277.

65. Chisaka,O. & Capecchi,M.R. 1991. Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox1.5. *Nature* 350, 473-479.

66. Chung,C.D., Liao,J., Liu,B., Rao,X., Jay,P., Berta,P. & Shuai,K. 1997. Specific inhibition of Stat3 signal transduction by PIAS3. *Science* 278, 1803-1805.

67. Conlon,F.L., Sedgwick,S.G., Weston,K.M. & Smith,J.C. 1996. Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. *Development* 122, 2427-2435.

68. Conway,S.J., Bundy,J., Chen,J., Dickman,E., Rogers,R. & Will,B.M. 2000. Decreased neural crest stem cell expansion is responsible for the conotruncal heart defects within the splotch (Sp(2H))/Pax3 mouse mutant. *Cardiovasc. Res.* 47, 314-328.

69. Conway,S.J., Godt,R.E., Hatcher,C.J., Leatherbury,L., Zolotouchnikov,V.V., Brotto,M.A., Copp,A.J., Kirby,M.L. & Creazzo,T.L. 1997a. Neural crest is involved in development of abnormal myocardial function. *J. Mol. Cell Cardiol.* 29, 2675-2685.

70. Conway,S.J., Henderson,D.J. & Copp,A.J. 1997b. Pax3 is required for cardiac neural crest migration in the mouse: evidence from the splotch (Sp2H) mutant. *Development* 124, 505-514.

71. Conway,S.J., Henderson,D.J., Kirby,M.L., Anderson,R.H. & Copp,A.J. 1997c. Development of a lethal congenital heart defect in the *splotch* (*Pax3*) mutant mouse. *Cardiovasc. Res.* 36, 163-173.

72. Cooper,S.T. & Hanson,I.M. 2005. A screen for proteins that interact with PAX6: C-terminal mutations disrupt interaction with HOMER3, DNCL1 and TRIM11. *BMC. Genet* 6, 43.

73. Corry,G.N. & Underhill,D.A. 2005. Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. *Pigment Cell Res.* 18, 427-438.

74. Cote,F., Boisvert,F.M., Grondin,B., Bazinet,M., Goodyer,C.G., Bazett-Jones,D.P. & Aubry,M. 2001. Alternative promoter usage and splicing of ZNF74 multifinger gene produce protein isoforms with a different repressor activity and nuclear partitioning. *DNA Cell Biol.* 20, 159-173.

75. Crepieux,P., Kwon,H., Leclerc,N., Spencer,W., Richard,S., Lin,R. & Hiscott,J. 1997. I kappaB alpha physically interacts with a cytoskeleton-associated protein through its signal response domain. *Mol. Cell Biol.* 17, 7375-7385.

76. Dai,M., Wang,P., Boyd,A.D., Kostov,G., Athey,B., Jones,E.G., Bunney,W.E., Myers,R.M., Speed,T.P., Akil,H., Watson,S.J. & Meng,F. 2005. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic acids researchX* 33, e175.

77. Daw,S.C.M., Taylor,C., Kraman,M., Call,K., Mao,J., Meitinger,T., Lipson,A., Goodship,J. & Scambler,P.J. 1996. A common region of 10p deleted in DiGeorge and velo-cardio-facial syndrome. *Nature Genetics* 13, 458-460.

78. de la Chapelle,A., Herva,R., Koivisto,M. & Aula,P. 1981. A deletion in chromosome 22 can cause DiGeorge syndrome. *Hum Genet* 57, 253-256.

79. De Lucia,F., Lorain,S., Scamps,C., Galisson,F., MacHold,J. & Lipinski,M. 2001. Subnuclear localization and mitotic phosphorylation of HIRA, the human homologue of Saccharomyces cerevisiae transcriptional regulators Hir1p/Hir2p. *Biochem J* 358, 447-455.

80. Dean,J.C., De Silva,D.C. & Reardon,W. 1998. Craniosynostosis and chromosome 22q11 deletion. *J. Med. Genet.* 35, 346.

81. DiGeorge,A.M. 1965. Discussion. *J. Pediatr.* 67, 907.

82. Digilio,M.C., Marino,B., Grazioli,S., Agostino,D., Giannotti,A. & Dallapiccola,B. 1996. Comparison of occurrence of genetic syndromes in ventricular septal defect with pulmonic stenosis (classic tetralogy of Fallot) versus ventricular septal defect with pulmonic atresia. *Am J Cardiol.* 77, 1375-1376.

83. Dodou,E., Verzi,M.P., Anderson,J.P., Xu,S.M. & Black,B.L. 2004. Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. *Development* 131, 3931-3942.

84. Dupont,E., Sansal,I., Evrard,C. & Rouget,P. 1998. Developmental pattern of expression of NPDC-1 and its interaction with E2F-1 suggest a role in the control of proliferation and differentiation of neural cells. *J. Neurosci. Res.* 51, 257-267.

85. Edelman,G.M. & Jones,F.S. 1998. Gene regulation of cell adhesion: a key step in neural morphogenesis. *Brain Res. Brain Res. Rev.* 26, 337-352.

86. Edelmann,L., Pandita,R., Spiteri,E., Funke,B., Goldberg,R., Palanisamy,N., Chaganti,R.S.K., Magenis,R.E., Shprintzen,R.J. & Morrow,B.E. 1999. A common molecular basis for rearrangement disorders on chromosome 22q11. *Human Molecular Genetics* 8, 1157-1167.

87. Edelmann,L., Stankiewicz,P., Spiteri,E., Pandita,R.K., Shaffer,L., Lupski,J.R. & Morrow,B.E. 2001. Two functional copies of the DGCR6 gene are present on human chromosome 22q11 due to a duplication of an ancestral locus. *Genome Research* 11, 208-217.

88. Edery,P., Attie,T., Amiel,J., Pelet,A., Eng,C., Hofstra,R.M., Martelli,H., Bidaud,C., Munnich,A. & Lyonnet,S. 1996. Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). *Nat. Genet.* 12, 442-444.

89. Edwards,D.R. & Denhardt,D.T. 1985. A study of mitochondrial and nuclear transcription with cloned cDNA probes. Changes in the relative abundance of mitochondrial transcripts after stimulation of quiescent mouse fibroblasts. *Exp. Cell Res.* 157, 127-143.

90. Emanuel,B.S., Budarf,B.S. & Scambler,P.J. 1998. The Genetic Basis of Conotruncal Heart Defects: The Chromosome 22q11.2 Deletion. In: *Heart Development* (Ed. by N.Rosenthal & R.Harvey), pp. 463-478. Academic Press.

91. Ensenauer,R.E., Adeyinka,A., Flynn,H.C., Michels,V.V., Lindor,N.M., Dawson,D.B., Thorland,E.C., Lorentz,C.P., Goldstein,J.L., McDonald,M.T., Smith,W.E., Simon-Fayard,E., Alexander,A.A., Kulharya,A.S., Ketterling,R.P., Clark,R.D. & Jalal,S.M. 2003. Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am. J. Hum. Genet.* 73, 1027-1040.

92. Epstein,D.J., Vekemans,M. & Gros,P. 1991. Splotch (Sp2H), a mutation affecting the development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax3. *Cell* 67, 767-774.

93. Epstein,E., Sela-Brown,A., Ringel,I., Kilav,R., King,S.M., Benashski,S.E., Yisraeli,J.K., Silver,J. & Naveh-Many,T. 2000a. Dynein light chain binding to a 3'-untranslated sequence mediates parathyroid hormone mRNA association with microtubules. *J. Clin. Invest* 105, 505-512.

94. Epstein,J., Cai,J., Glaser,T., Jepeal,L. & Maas,R. 1994. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. *Journal of Biological Chemistry* 269, 8355-8361.

95. Epstein,J.A., Li,J., Lang,D., Chen,F., Brown,C.B., Jin,F., Lu,M.M., Thomas,M., Liu,E., Wessels,A. & Lo,C.W. 2000b. Migration of cardiac neural crest cells in Splotch embryos. *Development* 127, 1869-1878.

96. Espindola,F.S., Suter,D.M., Partata,L.B., Cao,T., Wolenski,J.S., Cheney,R.E., King,S.M. & Mooseker,M.S. 2000. The light chain composition of chicken brain myosin-Va: calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN. *Cell Motil. Cytoskeleton* 47, 269-281.

97. Faedo,A., Ficara,F., Ghiani,M., Aiuti,A., Rubenstein,J.L. & Bulfone,A. 2002. Developmental expression of the T-box transcription factor T-bet/Tbx21 during mouse embryogenesis. *Mechanisms of Development* 116, 157-160.

98. Fang,J., Dagenais,S.L., Erickson,R.P., Arlt,M.F., Glynn,M.W., Gorski,J.L., Seaver,L.H. & Glover,T.W. 2000. Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome. *Am. J. Hum. Genet.* 67, 1382-1388.

99. Fang,X., Ji,H., Kim,S.W., Park,J.I., Vaught,T.G., Anastasiadis,P.Z., Ciesiolka,M. & McCrea,P.D. 2004. Vertebrate development requires ARVCF and p120 catenins and their interplay with RhoA and Rac. *JCB* 165, 87-98.

100. Farrell,M., Stadt,H., Wallis,K., Scambler,PJ., Hixon,R.L., Wolfe,R., Leatherbury,L. & Kirby,M. 1999. Persistent truncus arteriosus is associated with decreased expression of HIRA by cardiac neural crest cells in chick embryos. *Circulation Research* 84, 127-135.

101. Feijen,A., Goumans,M.J. & van den Eijnden-van Raaij AJ 1994. Expression of activin subunits, activin receptors and follistatin in postimplantation mouse embryos suggests specific developmental functions for different activins. *Development* 120, 3621-3637.

102. Feller,S.M. 2001. Crk family adaptors-signalling complex formation and biological roles. *Oncogene* 20, 6348-6371.

103. Fiegler,H., Carr,P., Douglas,E.J., Burford,D.C., Hunt,S., Scott,C.E., Smith,J., Vetrie,D., Gorman,P., Tomlinson,I.P. & Carter,N.P. 2003. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes. Cancer* 36, 361-374.

104. Fields,S. & Song,O.-K. 1989. A novel genetic system to detect protein-protein interactions. *Proc. Natl. Acd. sci USA* 340, 245-246.

105. Fleming,R.J., Gu,Y. & Hukriede,N.A. 1997. Serrate-mediated activation of Notch is specifically blocked by the product of the gene fringe in the dorsal compartment of the Drosophila wing imaginal disc. *Development* 124, 2973-2981.

106. Frank,D.U., Fotheringham,L.K., Brewer,J.A., Muglia,L.J., Tristani-Firouzi,M., Capecchi,M.R. & Moon,A.M. 2002. An Fgf8 mouse mutant phenocopies human 22q11 deletion syndrome. *Development* 129, 4591-4603.

107. Franz,T. 1989. Persistent truncus arteriosus in the Splotch mutant mouse. *Anat. Embryol.* 180, 457-464.

108. Franz,T. 1992. Neural tube defects without neural crest defects in splotch mice. *Teratology* 46, 599-604.

109. Franz,T. 1993. The Splotch (Sp1H) and Splotch-delayed (Spd) alleles: differential phenotypic effects on neural crest and limb musculature. *Anat. Embryol.* 187, 371-377.

110. Frascella,E., Toffolatti,L. & Rosolen,A. 1998. Normal and rearranged PAX3 expression in human rhabdomyosarcoma. *Cancer Genet. Cytogenet.* 102, 104-109.

111. Freedom,R.M., Rosen,F.S. & Nadas,A.S. 1972. Congenital cardiovascular disease and anomalies of the third and fourth pharyngeal pouch. *Circulation* 46, 165-172.

112. Friedlander,D.R., Milev,P., Karthikeyan,L., Margolis,R.K., Margolis,R.U. & Grumet,M. 1994. The neuronal chondroitin sulfate proteoglycan neurocan

binds to the neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits neuronal adhesion and neurite outgrowth. *JCB* 125, 669-680.

113. Fukushima,Y., Ohashi,H., Wakui,T., Nishida,T., Nakamura,Y., Hoshino,K., Ogawa,K. & Oh-ishi,T. DiGeorge syndrome with del(4)(q21.3q25): Possibility of the fourth chromosomal region responsible for DiGeorge syndrome. Am.J.Hum.Genet. 51, 306. 1992.

114. Funke,B., Puech,A., Saint-Jore,B., Pandita,R., Skoultchi,A. & Morrow,B. 1998. Isolation and characterization of a human gene containing a nuclear localization signal from the critical region for velo-cardio-facial syndrome on 22q11. *Genomics* 53, 146-154.

115. Gale,M., Jr., Blakely,C.M., Hopkins,D.A., Melville,M.W., Wambach,M., Romano,P.R. & Katze,M.G. 1998. Regulation of interferon-induced protein kinase PKR: modulation of P58IPK inhibitory function by a novel protein, P52rIPK. *Mol. Cell Biol.* 18, 859-871.

116. Galiana,E., Vernier,P., Dupont,E., Evrard,C. & Rouget,P. 1995. Identification of a neural-specific cDNA, NPDC-1, able to down-regulate cell proliferation and to suppress transformation. *Proc. Natl. Acad. Sci. U. S. A* 92, 1560-1564.

117. Galibert,M.D., Yavuzer,U., Dexter,T.J. & Goding,C.R. 1999. Pax3 and regulation of the melanocyte-specific tyrosinase-related protein-1 promoter. *J Biol. Chem.* 274, 26894-26900.

118. Galili,N., Epstein,J.A., Leconte,I., Nayak,S. & Buck,C.A. 1998. *Gscl*, a gene within the minimal DiGeorge critical region, is expressed in primordial germ cells and the developing pons. *Dev. Dyn* 212, 86-93.

119. Garcia-Minaur,S., Fantes,J., Murray,R.S., Porteous,M.E., Strain,L., Burns,J.E., Stephen,J. & Warner,J.P. 2002. A novel atypical 22q11.2 distal deletion in father and son. *J. Med. Genet.* 39, E62.

120. Garg,V., Kathiriya,I.S., Barnes,R., Schluterman,M.K., King,I.N., Butler,C.A., Rothrock,C.R., Eapen,R.S., Hirayama-Yamada,K., Joo,K., Matsuoka,R., Cohen,J.C. & Srivastava,D. 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424, 443-447.

121. Garg,V., Yamagishi,C., Hu,T., Kathiriya,I.S., Yamagishi,H. & Srivastava,D. 2001. Tbx1, a DiGeorge syndrome candidate gene, is regulated by sonic hedgehog during pharyngeal arch development. *Dev. Biol.* 235, 62-73.

122. Gehrmann,T. & Heilmeyer,L.M., Jr. 1998. Phosphatidylinositol 4-kinases. *European Journal of Biochemistry* 253, 357-370.

123. Geijsen,N., Uings,I.J., Pals,C., Armstrong,J., McKinnon,M., Raaijmakers,J.A., Lammers,J.W., Koenderman,L. & Coffer,P.J. 2001. Cytokine-specific transcriptional regulation through an IL-5Ralpha interacting protein. *Science* 293, 1136-1138.

124. Gennery,A.R., Barge,D., O'Sullivan,J.J., Flood,T.J., Abinun,M. & Cant,A.J. 2002. Antibody deficiency and autoimmunity in 22q11.2 deletion syndrome. *Arch. Dis. Child* 86, 422-425.

125. Gerber,H.P., Hillan,K.J., Ryan,A.M., Kowalski,J., Keller,G.A., Rangell,L., Wright,B.D., Radtke,F., Aguet,M. & Ferrara,N. 1999. VEGF is required for growth and survival in neonatal mice. *Development* 126, 1149-1159.

126. Ghosh,T.K., Packham,E.A., Bonser,A.J., Robinson,T.E., Cross,S.J. & Brook,J.D. 2001. Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Hum Mol. Genet* 10, 1983-1994.

127. Goding,C.R. T-box genes in development and disease. British Society for Developmental Biology . 2002.

128. Gogos,J.A., Santha,M., Takacs,Z., Beck,K.D., Luine,V., Lucas,L.R., Nadler,J.V. & Karayiorgou,M. 1999. The gene encoding proline dehydrogenase modulates sensorimotor gating in mice. *Nat Genet* 21, 434-439.

129. Goldberg,R., Motzkin,B., Marion,R., Scambler,P.J. & Shprintzen,R. 1993. Velo-Cardio-Facial Syndrome. *American Journal of Medical Genetics* 45, 313-319.

130. Goldmuntz,E., Driscoll,E., Budarf,M.L., Zackai,E.H., McDonald-McGinn,D.M., Biegel,J.A. & Emanuel,B.S. 1993. Microdeletion of chromosomal region 22q11 in patients with congenital conotruncal cardiac defects. *J. Med. Genet.* 30, 807-812.

131. Goldmuntz,E., Fedon,J., Roe,B. & Budarf,M.L. 1997. Molecular characterization of a serine/threonine kinase in the DiGeorge minimal critical region. *Gene* 198, 379-386.

132. Gong,L., Liu,M., Jen,J. & Yeh,E.T. 2000. GNB1L, a gene deleted in the critical region for DiGeorge syndrome on 22q11, encodes a G-protein beta-subunit-like polypeptide(1). *Biochim Biophys Acta* 1494, 185-188.

133. Gong,W., Emanuel,B.S., Collins,J., Kim,D.H., Wang,Z., Chen,F., Zhang,G., Roe,B. & Budarf,M.L. 1996. A transcription map of the DiGeorge and velo-cardio-facial syndrome critical region on 22q11. *Human Molecular Genetics* 5, 789-800.

134. Gong,W., Gottlieb,S., Collins,J., Blescia,A., Dietz,H., Goldmuntz,E., McDonald-McGinn,D.M., Zackai,E.H., Emanuel,B.S., Driscoll,D.A. & Budarf,M.L. 2001. Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects. *J. Med. Genet.* 38, E45.

135. Goodship,J., Cross,I., LiLing,J. & Wren,C. 1998. A population study of chromosome 22q11 deletions in infancy. *Arch. Dis. Child* 79, 348-351.

136. Goodship,J., Cross,I., Scambler,P. & Burn,J. 1995. Monozygotic twins with chromosome 22q11 deletion and discordant phenotype. *J. Med. Genet.* 32, 746-748.

137. Gothelf,D., Presburger,G., Levy,D., Nahmani,A., Burg,M., Berant,M., Blieden,L.C., Finkelstein,Y., Frisch,A., Apter,A. & Weizman,A. 2004. Genetic, developmental, and physical factors associated with attention deficit hyperactivity disorder in patients with velocardiofacial syndrome. *Am J Med. Genet B Neuropsychiatr. Genet* 126, 116-121.

138. Goulding,M., Lumsden,A. & Paquette,A.J. 1994. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development* 120, 957-971.

139. Goulding,M.D., Chalepakis,G., Deutsch,E., Erselius,J.R. & Gruss,P. 1991. Pax-3, a novel murine DNA binding protein expressed during early embryogenesis. *EMBO J.* 10, 1135-1147.

140. Graham,A. 2001. The development and evolution of the pharyngeal arches. *J. Anat.* 199, 133-141.

141. Greenberg,F., Courtney,K.B., Wessels,R.A., Huhta,J., Carpenter,R.J., Rich,D.C. & Ledbetter,D.H. 1988. Prenatal diagnosis of deletion 17p13 associated with DiGeorge anomaly. *Am. J. Med. Genet.* 31, 1-4.

142. Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N. & Shiekhattar,R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-240.

143. Grimm,C., Sporle,R., Schmid,T.E., Adler,I.D., Adamski,J., Schughart,K. & Graw,J. 1999. Isolation and embryonic expression of the novel mouse gene Hic1, the homologue of HIC1, a candidate gene for the Miller-Dieker syndrome. *Human Molecular Genetics* 8, 697-710.

144. Grootjans,J.J., Zimmermann,P., Reekmans,G., Smets,A., Degeest,G., Durr,J. & David,G. 1997. Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. *Proc. Natl. Acad. Sci. U. S. A* 94, 13683-13688.

145. Grossman,M.H., Emanuel,B.S. & Budarf,M.L. 1992. Chromosomal mapping of the human catechol-O-methyltransferase gene to 22q11.1----q11.2. *Genomics* 12, 822-825.

146. Gruber,P.J., Kubalak,S.W., Pexieder,T., Sucov,H.M., Evans,R.M. & Chien,K.R. 1996. RXR alpha deficiency confers genetic susceptibility for aortic sac, conotruncal, atrioventricular cushion, and ventricular muscle defects in mice. *J. Clin. Invest* 98, 1332-1343.

147. Guillaume,R. & Trudel,M. 2000. Distinct and common developmental expression patterns of the murine Pkd2 and Pkd1 genes. *Mechanisms of Development* 93, 179-183.

148. Guris,D.L., Duester,G., Papaioannou,V.E. & Imamoto,A. 2006. Dose-Dependent Interaction of Tbx1 and Crkl and Locally Aberrant RA Signaling in a Model of del22q11 Syndrome. *Dev. Cell* 10, 81-92.

149. Guris,D.L., Fantes,J., Tara,D., Druker,B.J. & Imamoto,A. 2001. Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. *Nat Genet* 27, 293-298.

150. Halford,S., Wadey,R., Roberts,C., Daw,S.C.M., Whiting,J.A., O'Donnell,H., Dunham,I., Bentley,D., Lindsay,E., Baldini,A., Francis,F., Lehrach,H., Williamson,R., Wilson,D.I., Goodship,J., Cross,I., Burn,J. & Scambler,P.J. 1993a. Isolation of a putative transcriptional regulator from the region of 22q11 deleted in DiGeorge Syndrome, Shprintzen syndrome and familial congenital heart disease. *Human Molecular Genetics* 2, 2099-2107.

151. Halford,S., Wilson,D.I., Daw,S.C.M., Roberts,C., Wadey,R., Kamath,S., Wickremasinghe,A., Burn,J., Goodship,J., Mattei,M.-G., Moorman,A.F.M. & Scambler,P.J. 1993b. Isolation of a gene expressed during early embryogenesis from the region of 22q11 commonly deleted in DiGeorge syndrome. *Human Molecular Genetics* 2, 1577-1582.

152. Hall,C., Nelson,D.M., Ye,X., Baker,K., DeCaprio,J.A., Seeholzer,S., Lipinski,M. & Adams,P.D. 2001. HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel cyclin-cdk2 substrate whose expression blocks S-phase progression. *Mol Cell Biol* 21, 1854-1865.

153. Hammond,P., Hutton,T.J., Allanson,J.E., Campbell,L.E., Hennekam,R.C., Holden,S., Patton,M.A., Shaw,A., Temple,I.K., Trotter,M., Murphy,K.C. & Winter,R.M. 2004. 3D analysis of facial morphology. *Am. J. Med. Genet. A* 126, 339-348.

154. Hanaoka,K., Qian,F., Boletta,A., Bhunia,A.K., Piontek,K., Tsiokas,L., Sukhatme,V.P., Guggino,W.B. & Germino,G.G. 2000. Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. *Nature* 408, 990-994.

155. Harrelson,Z., Kelly,R.G., Goldin,S.N., Gibson-Brown,J.J., Bollag,R.J., Silver,L.M. & Papaioannou,V.E. 2004. Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development. *Development* 131, 5041-5052.

156. Heathcote,K., Braybrook,C., Abushaban,L., Guy,M., Khetyar,M.E., Patton,M.A., Carter,N.D., Scambler,P.J. & Syrris,P. 2005. Common arterial trunk associated with a homeodomain mutation of NKX2.6. *Hum Mol. Genet* 14, 585-593.

157. Hedlund,E., Karsten,S.L., Kudo,L., Geschwind,D.H. & Carpenter,E.M. 2004. Identification of a Hoxd10-regulated transcriptional network and combinatorial interactions with Hoxa10 during spinal cord development. *J Neurosci. Res.* 75, 307-319.

158. Heisterkamp,N., Mulder,M.P., Langeveld,A., ten Hoeve,J., Wang,Z., Roe,B. & Groffen,J. 1995. Localization of the human mitochondrial citrate transporter protein gene to chromosome 22q11 in the DiGeorge syndrome critical region. *Genomics* 29, 451-456.

159. Heller,J.L. 1955. *CATCH22.*

160. Hemmati-Brivanlou,A., Kelly,O.G. & Melton,D.A. 1994. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* 77, 283-295.

161. Henderson,D.J., Conway,S.J. & Copp,A.J. 1999. Rib truncations and fusions in the Sp2H mouse reveal a role for Pax3 in specification of the ventro-lateral and posterior parts of the somite. *Dev. Biol.* 209, 143-158.

162. Henderson,D.J., Ybot-Gonzalez,P. & Copp,A.J. 1997. Over-expression of the chondroitin sulphate proteoglycan versican is associated with defective neural crest migration in the Pax3 mutant mouse (splotch). *Mechanisms of Development* 69, 39-51.

163. Herzig,R.P., Andersson,U. & Scarpulla,R.C. 2000. Dynein light chain interacts with NRF-1 and EWG, structurally and functionally related transcription factors from humans and drosophila. *J. Cell Sci.* 113 Pt 23, 4263-4273.

164. Hierck,B.P., Molin,D.G., Boot,M.J., Poelmann,R.E. & Gittenberger-de Groot,A.C. 2004. A chicken model for DGCR6 as a modifier gene in the DiGeorge critical region. *Pediatr. Res.* 56, 440-448.

165. Horb,M.E. & Thomsen,G.H. 1999. Tbx5 is essential for heart development. *Development* 126, 1739-1751.

166. Hu,T., Yamagishi,H., Maeda,J., McAnally,J., Yamagishi,C. & Srivastava,D. 2004. Tbx1 regulates fibroblast growth factors in the anterior heart field through a reinforcing autoregulatory loop involving forkhead transcription factors. *Development* 131, 5491-5502.

167. Hubank,M. & Schatz,D.G. 1994. Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic acids researchX* 22, 5640-5648.

168. Hunter,A.G., Dupont,B., McLaughlin,M., Hinton,L., Baker,E., Ades,L., Haan,E. & Schwartz,C.E. 2005. The Hunter-McAlpine syndrome results from duplication 5q35-qter. *Clin. Genet* 67, 53-60.

169. Husein,M., Chang,E., Cable,B., Karnell,M., Karnell,L.H. & Canady,J.W. 2004. Outcomes for children with submucous cleft palate and velopharyngeal insufficiency. *J Otolaryngol.* 33, 222-226.

170. Iemura,S., Yamamoto,T.S., Takagi,C., Uchiyama,H., Natsume,T., Shimasaki,S., Sugino,H. & Ueno,N. 1998. Direct binding of follistatin to a complex of bone-morphogenetic protein and its receptor inhibits ventral and epidermal cell fates in early Xenopus embryo. *Proc. Natl. Acad. Sci. U. S. A* 95, 9337-9342.

171. Iida,K., Koseki,H., Kakinuma,H., Kato,N., Mizutani-Koseki,Y., Ohuchi,H., Yoshioka,H., Noji,S., Kawamura,K., Kataoka,Y., Ueno,F., Taniguchi,M., Yoshida,N., Sugiyama,T. & Miura,N. 1997. Essential roles of the winged helix transcription factor MFH-1 in aortic arch patterning and skeletogenesis. *Development* 124, 4627-4638.

172. Ikeya,M., Lee,S.M., Johnson,J.E., McMahon,A.P. & Takada,S. 1997. Wnt signalling required for expansion of neural crest and CNS progenitors. *Nature* 389, 966-970.

173. Inoue,K., Tanabe,Y. & Lupski,J.R. 1999. Myelin deficiencies in both the central and the peripheral nervous systems associated with a SOX10 mutation. *Ann. Neurol.* 46, 313-318.

174. Ishii,J., Adachi,H., Aoki,J., Koizumi,H., Tomita,S., Suzuki,T., Tsujimoto,M., Inoue,K. & Arai,H. 2002. SREC-II, a new member of the scavenger receptor type F family, trans-interacts with SREC-I through its extracellular domain. *Journal of Biological Chemistry* 277, 39696-39702.

175. Ishkanian,A.S., Malloff,C.A., Watson,S.K., DeLeeuw,R.J., Chi,B., Coe,B.P., Snijders,A., Albertson,D.G., Pinkel,D., Marra,M.A., Ling,V., MacAulay,C. & Lam,W.L. 2004. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat. Genet* 36, 299-303.

176. Ivins,S., Lammerts,v.B., Roberts,C., James,C., Lindsay,E., Baldini,A., Ataliotis,P. & Scambler,P.J. 2005. Microarray analysis detects differentially expressed genes in the pharyngeal region of mice lacking Tbx1. *Dev. Biol.* 285, 554-569.

177. Iwarsson,E., Ahrlund-Richter,L., Inzunza,J., Fridstrom,M., Rosenlund,B., Hillensjo,T., Sjoblom,P., Nordenskjold,M. & Blennow,E. 1998. Preimplantation genetic diagnosis of DiGeorge syndrome. *Mol. Hum. Reprod.* 4, 871-875.

178. Jacquet,H., Demily,C., Houy,E., Hecketsweiler,B., Bou,J., Raux,G., Lerond,J., Allio,G., Haouzir,S., Tillaux,A., Bellegou,C., Fouldrin,G., Delamillieure,P., Menard,J.F., Dollfus,S., d'Amato,T., Petit,M., Thibaut,F., Frebourg,T. & Campion,D. 2004. Hyperprolinemia is a risk factor for schizoaffective disorder. *Mol. Psychiatry.*

179. Jacquet,H., Raux,G., Thibaut,F., Hecketsweiler,B., Houy,E., Demilly,C., Haouzir,S., Allio,G., Fouldrin,G., Drouin,V., Bou,J., Petit,M., Campion,D. & Frebourg,T. 2002. PRODH mutations and hyperprolinemia in a subset of schizophrenic patients. *Human Molecular Genetics* 11, 2243-2249.

180. Jaffrey,S.R. & Snyder,S.H. 1996. PIN: an associated protein inhibitor of neuronal nitric oxide synthase. *Science* 274, 774-777.

181. Jain,A.N., Tokuyasu,T.A., Snijders,A.M., Segraves,R., Albertson,D.G. & Pinkel,D. 2002. Fully automatic quantification of microarray image data. *Genome Research* 12, 325-332.

182. James,P., Halladay,J. & Craig,E.A. 1996. Genomic libraries and a host strain designed for efficient two-hybrid selection in yeast. *Genetics* 144, 1425-1436.

183. Jawad,A.F., McDonald-McGinn,D.M., Zackai,E. & Sullivan,K.E. 2001. Immunologic features of chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome). *J. Pediatr.* 139, 715-723.

184. Jerome,L.A. & Papaioannou,V.E. 2001. DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. *Nat. Genet.* 27, 286-291.

185. Kaiser,F.J., Tavassoli,K., Van Den Bemd,G.J., Chang,G.T., Horsthemke,B., Moroy,T. & Ludecke,H.J. 2003. Nuclear interaction of the dynein light chain

LC8a with the TRPS1 transcription factor suppresses the transcriptional repression activity of TRPS1. *Human Molecular Genetics* 12, 1349-1358.

186. Kallioniemi,A., Kallioniemi,O.P., Sudar,D., Rutovitz,D., Gray,J.W., Waldman,F. & Pinkel,D. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818-821.

187. Kariyazono,H., Ohno,T., Ihara,K., Igarashi,H., Joh-o K, Ishikawa,S. & Hara,T. 2001. Rapid detection of the 22q11.2 deletion with quantitative real-time PCR. *Mol. Cell Probes* 15, 71-73.

188. Kawame,H., Adachi,M., Tachibana,K., Kurosawa,K., Ito,F., Gleason,M.M., Weinzimer,S., Levitt-Katz,L., Sullivan,K. & McDonald-McGinn,D.M. 2001. Graves' disease in patients with 22q11.2 deletion. *J. Pediatr.* 139, 892-895.

189. Kennan,A., Aherne,A., Palfi,A., Humphries,M., McKee,A., Stitt,A., Simpson,D.A., Demtroder,K., Orntoft,T., Ayuso,C., Kenna,P.F., Farrar,G.J. & Humphries,P. 2002. Identification of an IMPDH1 mutation in autosomal dominant retinitis pigmentosa (RP10) revealed following comparative microarray analysis of transcripts derived from retinas of wild-type and Rho(-/-) mice. *Hum Mol. Genet* 11, 547-557.

190. Khan,J., Bittner,M.L., Saal,L.H., Teichmann,U., Azorsa,D.O., Gooden,G.C., Pavan,W.J., Trent,J.M. & Meltzer,P.S. 1999. cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene [In Process Citation]. *Proc Natl Acad Sci U S A* 96, 13264-13269.

191. Kim,E., Arnould,T., Sellin,L.K., Benzing,T., Fan,M.J., Gruning,W., Sokol,S.Y., Drummond,I. & Walz,G. 1999. The polycystic kidney disease 1 gene product modulates Wnt signaling. *Journal of Biological Chemistry* 274, 4947-4953.

192. Kinouchi,A., Mori,K., Ando,M. & Takao,A. Facial appearance of patients with conotruncal anomalies. 17[Pediatrics Japan], 84. 1976.

193. Kioussi,C., Gross,M.K. & Gruss,P. 1995. Pax3: a paired domain gene as a regulator in PNS myelination. *Neuron* 15, 553-562.

194. Kirby,M.L., Gale,T.F. & Stewart,D.E. 1983. Neural crest cells contribute to normal aorticopulmonary septation. *Science* 220, 1059-1061.

195. Kirchhoff,M., Rose,H. & Lundsteen,C. 2001. High resolution comparative genomic hybridisation in clinical cytogenetics. *J. Med. Genet.* 38, 740-744.

196. Kleinjan,D.A., Seawright,A., Schedl,A., Quinlan,R.A., Danes,S. & van,H., V 2001. Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol. Genet* 10, 2049-2059.

197. Knoll,J.H., Asamoah,A., Pletcher,B.A. & Wagstaff,J. 1995. Interstitial duplication of proximal 22q: phenotypic overlap with cat eye syndrome. *Am. J. Med. Genet.* 55, 221-224.

198. Koblar,S.A., Murphy,M., Barrett,G.L., Underhill,A., Gros,P. & Bartlett,P.F. 1999. Pax-3 regulates neurogenesis in neural crest-derived precursor cells. *J. Neurosci. Res.* 56, 518-530.

199. Kochilas,L.K., Merscher-Gomez,S., Lu,M.M., Potluri,V., Liao,J., Kucherlapati,R., Morrow,B. & Epstein,J.A. 2002. The role of neural crest during cardiac development in a mouse model of DiGeorge syndrome. *Dev. Biol.* 251, 157-166.

200. Krause,A., Zacharias,W., Camarata,T., Linkhart,B., Law,E., Lischke,A., Miljan,E. & Simon,H.G. 2004. Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Dev. Biol.* 273, 106-120.

201. Kruger,O., Plum,A., Kim,J.S., Winterhager,E., Maxeiner,S., Hallas,G., Kirchhoff,S., Traub,O., Lamers,W.H. & Willecke,K. 2000. Defective vascular development in connexin 45-deficient mice. *Development* 127, 4179-4193.

202. Kumai,M., Nishii,K., Nakamura,K., Takeda,N., Suzuki,M. & Shibata,Y. 2000. Loss of connexin45 causes a cushion defect in early cardiogenesis. *Development* 127, 3501-3512.

203. Kurahashi,H., Nakayama,T., Osugi,Y., Tsuda,E., Masuno,M., Imaizumi,K., Kamiya,T., Sano,T., Okada,S. & Nishisho,I. 1996. Deletion mapping of 22q11 in CATCH22 syndrome: identification of a second critical region. *Am. J. Hum. Genet.* 58, 1377-1881.

204. Kwang,S.J., Brugger,S.M., Lazik,A., Merrill,A.E., Wu,L.Y., Liu,Y.H., Ishii,M., Sangiorgi,F.O., Rauchman,M., Sucov,H.M., Maas,R.L. & Maxson,R.E., Jr. 2002. Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest. *Development* 129, 527-538.

205. Lachman,H.M., Morrow,B., Shprintzen,R., Veit,S., Parsia,S.S., Faedda,G., Goldberg,R., Kucherlapati,R. & Papolos,D.F. 1996. Association of codon 108/158 catechol-O-methyltransferase gene polymorphism with the psychiatric manifestations of velocardiofacial syndrome. *Am. J. Med. Genet.* 67, 468-472.

206. Lalani,S.R., Safiullah,A.M., Molinari,L.M., Fernbach,S.D., Martin,D.M. & Belmont,J.W. 2004. SEMA3E mutation in a patient with CHARGE syndrome. *J Med. Genet* 41, e94.

207. Lammer,E.J., Chen,D.T., Hoar,N.D., Agnish,P.J., Benke,J.T., Braun,C.J., Curry,P.M., Fernhoff,A.W., Grix,A.T., Lott,J.M., Richard,J.M. & Sun,S.C. 1986. Retinoic acid embryopathy. A new human teratogen and a mechanistic hypothesis. *New England Journal of Medicne* 313, 837-841.

208. Lamolet,B., Pulichino,A., Lamonerie,T., Gauthier,Y., Brue,T., Enjalbert,A. & Drouin,J. 2001. A pituitary cell-restricted t box factor, tpit, activates pomc transcription in cooperation with pitx homeoproteins. *Cell* 104, 849-859.

209. Lang,D. & Epstein,J.A. 2003. Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. *Hum Mol. Genet* 12, 937-945.

210. Law,D.J., Garvey,N., Agulnik,S.I., Perlroth,V., Hahn,O.M., Rhinehart,R.E., Gebuhr,T.C. & Silver,L.M. 1998. TBX10, a member of the Tbx1-subfamily of conserved developmental genes, is located at human chromosome 11q13 and proximal mouse chromosome 19. *Mamm Genome* 9, 397-399.

211. LeBlanc,E., Gereau,S. & Bassila,M. Prevalence of laryngeal anomalies in velocardiofacial syndrome. Presented at the American Cleft Palate Craniofacial Association . 1996.

212. Lee,H.J., Adham,I.M., Schwarz,G., Kneussel,M., Sass,J.O., Engel,W. & Reiss,J. 2002. Molybdenum cofactor-deficient mice resemble the phenotype of human patients. *Human Molecular Genetics* 11, 3309-3317.

213. Lee,S.H., Schloss,D.J., Jarvis,L., Krasnow,M.A. & Swain,J.L. 2001. Inhibition of angiogenesis by a mouse sprouty protein. *Journal of Biological Chemistry* 276, 4128-4133.

214. Lehmann,R. & Tautz,D. 1994. In situ hybridization to RNA. *Methods Cell Biol.* 44, 575-598.

215. Levy,A., Michel,G., Lemerrer,M. & Philip,N. 1997. Idiopathic thrombocytopenic purpura in two mothers of children with DiGeorge sequence. *Am. J. Med. Genet.* 69, 356-359.

216. Levy,C., Nechushtan,H. & Razin,E. 2002. A new role for the STAT3 inhibitor, PIAS3: a repressor of microphthalmia transcription factor. *Journal of Biological Chemistry* 277, 1962-1966.

217. Li,J., Liu,K.C., Jin,F., Lu,M.M. & Epstein,J.A. 1999. Transgenic rescue of congenital heart disease and spina bifida in Splotch mice. *Development* 126, 2495-2503.

218. Li,L., Krantz,I.D., Deng,Y., Genin,A., Banta,A.B., Collins,C.C., Qi,M., Trask,B.J., Kuo,W.L., Cochran,J., Costa,T., Pierpont,M.E., Rand,E.B., Piccoli,D.A., Hood,L. & Spinner,N.B. 1997a. Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1 [see comments]. *Nat Genet* 16, 243-251.

219. Li,Q.Y., Newbury-Ecob,R.A., Terrett,J.A., Wilson,D.I., Curtis,A.R.J., Yi,C.H., Gebuhr,T., Bullen,P.J., Robson,S.C., Strachan,T., Bonnet,D., Lyonnet,S., Young,I.D., Raeburn,A., Buckler,A.J., Law,D.J. & Brook,J.D. 1997b. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nature Genetics* 15, 21-29.

220. Liao,J., Kochilas,L., Nowotschin,S., Arnold,J.S., Aggarwal,V.S., Epstein,J.A., Brown,M.C., Adams,J. & Morrow,B.E. 2004. Full spectrum of malformations in velo-cardio-facial syndrome/DiGeorge syndrome mouse models by altering Tbx1 dosage. *Human Molecular Genetics* 13, 1577-1585.

221. Lin,L., Bu,L., Cai,C.L., Zhang,X. & Evans,S. 2006. Isl1 is upstream of sonic hedgehog in a pathway required for cardiac morphogenesis. *Dev. Biol.*

222. Lindsay,E.A. & Baldini,A. 1997. A mouse gene (Dgcr6) related to the Drosophila gonadal gene is expressed in early embryogenesis and is the

homolog of a human gene deleted in DiGeorge syndrome. *Cytogenet. Cell Genet.* 79, 243-247.

223. Lindsay,E.A. & Baldini,A. 2001. Recovery from arterial growth delay reduces penetrance of cardiovascular defects in mice deleted for the DiGeorge syndrome region. *Hum Mol Genet* 10, 997-1002.

224. Lindsay,E.A., Botta,A., Jurecic,V., Carattini-Rivera,S., Cheah,Y.-C., Rosenblatt,H.M., Bradley,A. & Baldini,A. 1999a. Congenital heart disease in mice deficient for the DiGeorge syndrome region. *Nature* 401, 379-383.

225. Lindsay,E.A., Morris,M.A., Gos,A., Nestadt,G., Wolyniec,P.S., Lasseter,V.K., Shprintzen,R., Antonarakis,S.E., Baldini,A. & Pulver,A.E. 1995. Schizophrenia and chromosomal deletions within 22q11.2. *Am. J. Hum. Genet.* 56, 1502-1503.

226. Lindsay,E.A., Su,D. & Baldini,A. Embryonic recovery is the mechanism for reduced penetrance in mice deficient for the DiGeorge syndrome region. Am.J.Hum.Genet. 65, A33. 1999b.

227. Lindsay,E.A., Vitelli,F., Su,H., Morishima,M., Huynh,T., Pramparo,T., Jurecic,V., Ogunrinu,G., Sutherland,H.F., Scambler,P.J., Bradley,A. & Baldini,A. 2001. *Tbx1* haploinsufficiency identified by functional scanning of the DiGeorge syndrome region is the cause of aortic arch defects in mice. *Nature* 401, 97-101.

228. Liu,J., Wang,X.B., Park,D.S. & Lisanti,M.P. 2002. Caveolin-1 expression enhances endothelial capillary tubule formation. *Journal of Biological Chemistry* 277, 10661-10668.

229. Liu,W., Selever,J., Wang,D., Lu,M.F., Moses,K.A., Schwartz,R.J. & Martin,J.F. 2004. Bmp4 signaling is required for outflow-tract septation and branchial-arch artery remodeling. *Proc. Natl. Acad. Sci. U. S. A* 101, 4489-4494.

230. Livak,K.J. & Schmittgen,T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.

231. Lo,K.W., Kan,H.M., Chan,L.N., Xu,W.G., Wang,K.P., Wu,Z., Sheng,M. & Zhang,M. 2005. The 8-kDa dynein light chain binds to p53-binding protein 1 and mediates DNA damage-induced p53 nuclear accumulation. *J Biol. Chem.* 280, 8172-8179.

232. Lobdell,D.H. 1959. Congenital absence of the parathyroid glands. *Archives of Path* 68, 412-415.

233. Lohnes,D., Mark,M., Mendelsohn,C., Dolle,P., Dierich,A., Gorry,P., Gansmuller,A. & Chambon,P. 1994. Function of the retinoic acid receptors (RARs) during development. (I) Craniofacial and skeletal abnormalities in RAR double mutants. *Development* 120, 2723-2748.

234. Loppin,B., Bonnefoy,E., Anselme,C., Laurencon,A., Karr,T.L. & Couble,P. 2005. The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* 437, 1386-1390.

235. Lorain,S., Quivy,J.-P., Monier-Gavelle,F., Scamps,C., Lecluse,Y., Almouzni,G. & Lipinski,M. 1998. Core histones and HIRIP3, a novel histone-binding protein, directly interact with the WD repeat protein HIRA. *Molecular and Cellular Biology* 18, 5546-5556.

236. Maalouf,N.M., Sakhaee,K. & Odvina,C.V. 2004. A case of chromosome 22q11 deletion syndrome diagnosed in a 32-year-old man with hypoparathyroidism. *J. Clin. Endocrinol. Metab* 89, 4817-4820.

237. Magnaghi,P., Roberts,C., Lorain,S., Lipinski,M. & Scambler,P.J. 1998. HIRA, a mammalian homologue of *Saccharomyces cerevisiae* transcriptional co-repressors, interacts with Pax3. *Nature Genetics* 20, 74-77.

238. Malakhov,M.P., Malakhova,O.A., Kim,K.I., Ritchie,K.J. & Zhang,D.E. 2002. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *Journal of Biological Chemistry* 277, 9976-9981.

239. Manley,N. & Capecchi,M.R. 1995. The role of Hoxa-3 in mouse thymus and thyroid development. *Development* 121, 1989-2003.

240. Margue,C.M., Bernasconi,M., Barr,F.G. & Schafer,B.W. 2000. Transcriptional modulation of the anti-apoptotic protein BCL-XL by the paired box transcription factors PAX3 and PAX3/FKHR. *Oncogene* 19, 2921-2929.

241. Maroto,M., Reshef,R., Munsterberg,A.E., Koester,S., Goulding,M. & Lassar,A.B. 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell* 89, 139-148.

242. Marvin,M.J., Di Rocco,G., Gardiner,A., Bush,S.M. & Lassar,A.B. 2001. Inhibition of Wnt activity induces heart formation from posterior mesoderm. *Genes Dev.* 15, 316-327.

243. Matsuoka,R., Kimua,M., Scambler,P.J., Morrow,B.E., Imamura,S., Minoshima,S., Shimizu,N., Yamagish,H., Joh-o,K., Watanabe,S., Oyama,K., Saji,T., Ando,M., Takao,A. & Momma,K. 1998. Molecular and clinical study of 183 patients with conotruncal anomaly face. *Hum. Genet.* 103, 70-80.

244. Matzuk,M.M., Lu,N., Vogel,H., Sellheyer,K., Roop,D.R. & Bradley,A. 1995. Multiple defects and perinatal death in mice deficient in follistatin. *Nature* 374, 360-363.

245. Mayanil,C.S., George,D., Freilich,L., Miljan,E.J., Mania-Farnell,B., McLone,D.G. & Bremer,E.G. 2001. Microarray analysis detects novel Pax3 downstream target genes. *Journal of Biological Chemistry* 276, 49299-49309.

246. Mayanil,C.S., Pool,A., Nakazaki,H., Reddy,A., Farnell,B.M., Yun,B., George,D., McLone,D.G. & Bremer,E.G. 2006. Regulation of murine Tgfbeta 2 by Pax3 during early embryonic development. *J Biol. Chem.*

247. Maynard,T.M., Haskell,G.T., Bhasin,N., Lee,J.M., Gassman,A.A., Lieberman,J.A. & LaMantia,A.S. 2002. RanBP1, a velocardiofacial/DiGeorge syndrome candidate gene, is expressed at sites of mesenchymal/epithelial induction. *Mechanisms of Development* 111, 177-180.

248. Maynard,T.M., Haskell,G.T., Peters,A.Z., Sikich,L., Lieberman,J.A. & LaMantia,A.S. 2003. A comprehensive analysis of 22q11 gene expression in the developing and adult brain. *Proc. Natl. Acad. Sci. U. S. A* 100, 14433-14438.

249. McDermid,H.E. & Morrow,B.E. 2002. Genomic disorders on 22q11. *Am. J. Hum. Genet.* 70, 1077-1088.

250. McDonald-McGinn,D.M., Kirschner,R., Goldmuntz,E., Sullivan,K., Eicher,P., Gerdes,M., Moss,E., Solot,C., Wang,P., Jacobs,I., Handler,S., Knightly,C., Heher,K., Wilson,M., Ming,J.E., Grace,K., Driscoll,D., Pasquariello,P., Randall,P., LaRossa,D., Emanuel,B.S. & Zackai,E.H. 1999. The Philadelphia story: the 22q11.2 deletion: report on 250 patients. *Genet Couns.* 10, 11-24.

251. McDonald-McGinn,D.M., Tonnesen,M.K., Laufer-Cahana,A., Finucane,B., Driscoll,D.A., Emanuel,B.S. & Zackai,E.H. 2001. Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: cast a wide FISHing net! *Genet. Med.* 3, 23-29.

252. McElhinney,D.B., Clark,B.J., III, Weinberg,P.M., Kenton,M.L., McDonald-McGinn,D., Driscoll,D.A., Zackai,E.H. & Goldmuntz,E. 2001. Association of chromosome 22q11 deletion with isolated anomalies of aortic arch laterality and branching. *J Am Coll. Cardiol.* 37, 2114-2119.

253. McElhinney,D.B., Driscoll,D.A., Emanuel,B.S. & Goldmuntz,E. 2003a. Chromosome 22q11 deletion in patients with truncus arteriosus. *Pediatr. Cardiol.* 24, 569-573.

254. McElhinney,D.B., Driscoll,D.A., Levin,E.R., Jawad,A.F., Emanuel,B.S. & Goldmuntz,E. 2003b. Chromosome 22q11 deletion in patients with ventricular septal defect: frequency and associated cardiovascular anomalies. *Pediatrics* 112, e472.

255. McKerracher,L. & Winton,M.J. 2002. Nogo on the go. *Neuron* 36, 345-348.

256. McKie,J.M., Sutherland,H.F., Harvey,E., Kim,U.-J. & Scambler,P.J. 1998. A human gene similar to Drosophila melanogaster peanut maps to the DiGeorge Syndrome region of 22q11. *Hum. Genet.* 101, 6-12.

257. McQuade,L., Christodoulou,J., Budarf,B., Sachdev,R., Wilson,M., Emanuel,B. & Colley,A. 1999. Patient with a 22q11.2 deletion with no overlap of the minimal DiGeorge syndrome critical region (MDGCR). *Am. J. Med. Genet.* 86, 27-33.

258. Merscher,S., Funke,B., Epstein,J.A., Heyer,J., Puech,A., Lu,M.M., Xavier,R.J., Demay,M.B., Russell,R.G., Factor,S., Tokooya,K., St.Jore,B., Lopez,M., Pandita,R.K., Lia,M., Carrion,D., Schorle,H., Kobler,J.R., Scambler,P.J., Wynshaw-Boris,A., Skoultchi,A., Morrow,B.E. & Kucherlapati,R. 2001. *TBX1* is responsible for the cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* 104, 619-629.

259. Meyers,E.N., Lewandoski,M. & Martin,G.R. 1998. An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. *Nat Genet* 18, 136-141.

260. Mitnick,R.J., Bello,J.A. & Shprintzen,R.J. 1994. Brain anomalies in velo-cardio-facial syndrome. *Am. J. Med. Genet.* 54, 100-106.

261. Miyamori,H., Takino,T., Kobayashi,Y., Tokai,H., Itoh,Y., Seiki,M. & Sato,H. 2001. Claudin promotes activation of pro-matrix metalloproteinase-2 mediated by membrane-type matrix metalloproteinases. *Journal of Biological Chemistry* 276, 28204-28211.

262. Moase,C.E. & Trasler,D.G. 1990. Delayed neural crest emigration from Sp and Sp^d mouse neural tube explants. *Teratology* 42, 171-182.

263. Momeni,P., Glockner,G., Schmidt,O., von Holtum,D., Albrecht,B., Gillessen-Kaesbach,G., Hennekam,R., Meinecke,P., Zabel,B., Rosenthal,A., Horsthemke,B. & Ludecke,H.J. 2000. Mutations in a new gene, encoding a zinc-finger protein, cause tricho-rhino-phalangeal syndrome type I. *Nat. Genet* 24, 71-74.

264. Momma,K., Kondo,C., Matsuoka,R. & Takao,A. 1996. Cardiac anomalies associated with a chromosome 22q11 deletion in patients with conotruncal anomaly face syndrome. *Am J Cardiol.* 78, 591-594.

265. Monni,O., Barlund,M., Mousses,S., Kononen,J., Sauter,G., Heiskanen,M., Paavola,P., Avela,K., Chen,Y., Bittner,M.L. & Kallioniemi,A. 2001. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. *Proc. Natl. Acad. Sci. U. S. A* 98, 5711-5716.

266. Morrow,B., Goldberg,R., Carlson,C., Das Gupta,R., Sirotkin,H., Collins,J., Dunham,I., O'Donnell,H., Scambler,P., Shprintzen,R. & Kucherlapati,R. 1995. Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am. J. Hum. Genet.* 56, 1379-1390.

267. Mukai,J., Liu,H., Burt,R.A., Swor,D.E., Lai,W.S., Karayiorgou,M. & Gogos,J.A. 2004. Evidence that the gene encoding ZDHHC8 contributes to the risk of schizophrenia. *Nat. Genet.* 36, 725-731.

268. Mulder,G.B., Manley,N., Grant,J., Schmidt,K., Zeng,W., Eckhoff,C. & Maggio-Price,L. 2000. Effects of excess vitamin A on development of cranial neural crest-derived structures: a neonatal and embryologic study. *Teratology* 62, 214-226.

269. Mulder,G.B., Manley,N. & Maggio-Price,L. 1998. Retinoic acid-induced thymic abnormalities in the mouse are associated with altered pharyngeal morphology, thymocyte maturation defects, and altered expression of Hoxa3 and Pax1. *Teratology* 58, 263-275.

270. Murphy,K.C. 2002. Schizophrenia and velo-cardio-facial syndrome. *Lancet* 359, 426-430.

271. Murphy,K.C., Jones,L.A. & Owen,M.J. 1999. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 56, 940-945.

272. Muto,S., Aiba,A., Saito,Y., Nakao,K., Nakamura,K., Tomita,K., Kitamura,T., Kurabayashi,M., Nagai,R., Higashihara,E., Harris,P.C., Katsuki,M. & Horie,S.

2002. Pioglitazone improves the phenotype and molecular defects of a targeted Pkd1 mutant. *Human Molecular Genetics* 11, 1731-1742.

273. Nadon,R. & Shoemaker,J. 2002. Statistical issues with microarrays: processing and analysis. *Trends Genet* 18, 265-271.

274. Naiche,L.A. & Papaioannou,V.E. 2003. Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. *Development* 130, 2681-2693.

275. Nakamura,T., Takio,K., Eto,Y., Shibai,H., Titani,K. & Sugino,H. 1990. Activin-binding protein from rat ovary is follistatin. *Science* 247, 836-838.

276. Niebuhr,E. 1978. Cytologic observations in 35 individuals with a 5p- karyotype. *Hum Genet* 42, 143-156.

277. Niederreither,K., Vermot,J., Messaddeq,N., Schuhbaur,B., Chambon,P. & Dolle,P. 2001. Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. *Development* 128, 1019-1031.

278. Nilsen,H., Stamp,G., Andersen,S., Hrivnak,G., Krokan,H.E., Lindahl,T. & Barnes,D.E. 2003. Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene* 22, 5381-5386.

279. Novelli,G., Amati,F. & Dallapiccola,B. 2000. Individual haploinsufficient loci and the complex phenotype of DiGeorge syndrome. *Mol. Med. Today* 6, 10-11.

280. Nowotschin,S., Liao,J., Gage,P.J., Epstein,J.A., Campione,M. & Morrow,B.E. 2006. Tbx1 affects asymmetric cardiac morphogenesis by regulating Pitx2 in the secondary heart field. *Development* 133, 1565-1573.

281. O'Donnell,H., McKeown,C., Gould,C., Morrow,B. & Scambler,P. 1997. Detection of a deletion within 22q11 which has no overlap with the DiGeorge syndrome critical region. *Am. J. Hum. Genet.* 60, 1544-1548.

282. Olson,J.A., Jr. 2004. Application of microarray profiling to clinical trials in cancer. *Surgery* 136, 519-523.

283. Pang,Q., Fagerlie,S., Christianson,T.A., Keeble,W., Faulkner,G., Diaz,J., Rathbun,R.K. & Bagby,G.C. 2000. The Fanconi anemia protein FANCC binds to and facilitates the activation of STAT1 by gamma interferon and hematopoietic growth factors. *Mol. Cell Biol.* 20, 4724-4735.

284. Pani,L., Horal,M. & Loeken,M.R. 2002. Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. *Genes Dev.* 16, 676-680.

285. Panin,V.M., Papayannopoulos,V., Wilson,R. & Irvine,K.D. 1997. Fringe modulates Notch-ligand interactions. *Nature* 387, 908-912.

286. Paplos,D.F., Faedda,G.L., Veit,S., Goldberg,R., Morrow,B., Kucherlapati,R. & Shprintzen,R.J. 1996. Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome: does a hemizygous deletion of chromosome 22q11

result in bipolar affective disorder? *Ameriacan Journal of Psychiatry* 153, 1541-1547.

287. Parwani,A., Duncan,E.J., Bartlett,E., Madonick,S.H., Efferen,T.R., Rajan,R., Sanfilipo,M., Chappell,P.B., Chakravorty,S., Gonzenbach,S., Ko,G.N. & Rotrosen,J.P. 2000. Impaired prepulse inhibition of acoustic startle in schizophrenia. *Biol. Psychiatry* 47, 662-669.

288. Paylor,R., Ataliotis,P., Glaser,B., Mupo,A., Spencer,C., Sobotka,A., Sparks,C., Choi,C.-H., Oghalai,J., Curran,S., Murphy,K., Williams,N., O'Donovan.M.C, Owen,M.J., Scambler,P.J. & Lindsay,E. Tbx1 haploinsuffiency is linked to behavioural disorders in mice and humans: Implications for 22q11 deletion syndrome. Proc.Natl.Acad.Sci . 2006.

289. Paylor,R., McIlwain,K.L., McAninch,R., Nellis,A., Yuva-Paylor,L.A., Baldini,A. & Lindsay,E.A. 2001. Mice deleted for the DiGeorge/velocardiofacial syndrome region show abnormal sensorimotor gating and learning and memory impairments. *Human Molecular Genetics* 10, 2645-2650.

290. Peifer,M. & Polakis,P. 2000. Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. *Science* 287, 1606-1609.

291. Persson,C., Lohmander,A., Jonsson,R., Oskarsdottir,S. & Soderpalm,E. 2003. A prospective cross-sectional study of speech in patients with the 22q11 deletion syndrome. *J Commun. Disord.* 36, 13-47.

292. Peters,K.G., De Vries,C. & Williams,L.T. 1993. Vascular endothelial growth factor receptor expression during embryogenesis and tissue repair suggests a role in endothelial differentiation and blood vessel growth. *Proc. Natl. Acad. Sci. U. S. A* 90, 8915-8919.

293. Pfeifer,D., Kist,R., Dewar,K., Devon,K., Lander,E.S., Birren,B., Korniszewski,L., Back,E. & Scherer,G. 1999. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet* 65, 111-124.

294. Phelan,S.A., Ito,M. & Loeken,M.R. 1997. Neural tube defects in embryos of diabetic mice: role of the Pax-3 gene and apoptosis. *Diabetes* 46, 1189-1197.

295. Piccolo,S., Sasai,Y., Lu,B. & De Robertis,E.M. 1996. Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* 86, 589-598.

296. Pinkel,D., Landegent,J., Collins,C., Fuscoe,J., Segraves,R., Lucas,J. & Gray,J. 1988. Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. U. S. A* 85, 9138-9142.

297. Pinkel,D., Segraves,R., Sudar,D., Clark,S., Poole,I., Kowbel,D., Collins,C., Kuo,W.L., Chen,C., Zhai,Y., Dairkee,S.H., Ljung,B.M., Gray,J.W. & Albertson,D.G. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.* 20, 207-211.

298. Plageman,T.F., Jr. & Yutzey,K.E. 2004. Differential expression and function of Tbx5 and Tbx20 in cardiac development. *Journal of Biological Chemistry* 279, 19026-19034.

299. Pollack,J.R., Sorlie,T., Perou,C.M., Rees,C.A., Jeffrey,S.S., Lonning,P.E., Tibshirani,R., Botstein,D., Borresen-Dale,A.L. & Brown,P.O. 2002. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc. Natl. Acad. Sci. U. S. A* 99, 12963-12968.

300. Prescott,K., Ivins,S., Hubank,M., Lindsay,E., Baldini,A. & Scambler,P. 2005a. Microarray analysis of the Df1 mouse model of the 22q11 deletion syndrome. *Hum Genet* 116, 486-496.

301. Prescott,K., Woodfine,K., Stubbs,P., Super,M., Kerr,B., Palmer,R., Carter,N.P. & Scambler,P. 2005b. A novel 5q11.2 deletion detected by microarray comparative genomic hybridisation in a child referred as a case of suspected 22q11 deletion syndrome. *Hum Genet* 116, 83-90.

302. Prieve,M.G. & Moon,R.T. 2003. Stromelysin-1 and mesothelin are differentially regulated by Wnt-5a and Wnt-1 in C57mg mouse mammary epithelial cells. *BMC. Dev. Biol.* 3, 2.

303. Prols,F., Mayer,M.P., Renner,O., Czarnecki,P.G., Ast,M., Gassler,C., Wilting,J., Kurz,H. & Christ,B. 2001. Upregulation of the cochaperone Mdg1 in endothelial cells is induced by stress and during in vitro angiogenesis. *Exp. Cell Res.* 269, 42-53.

304. Puffenberger,E.G., Hosoda,K., Washington,S.S., Nakao,K., deWit,D., Yanagisawa,M. & Chakravart,A. 1994. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. *Cell* 79, 1257-1266.

305. Rape,M., Hoppe,T., Gorr,I., Kalocay,M., Richly,H. & Jentsch,S. 2001. Mobilization of processed, membrane-tethered SPT23 transcription factor by CDC48(UFD1/NPL4), a ubiquitin-selective chaperone. *Cell* 107, 667-677.

306. Rasmussen,S.A., Williams,C.A., Ayoub,E.M., Sleasman,J.W., Gray,B.A., Bent-Williams,A., Stalker,H.J. & Zori,R.T. 1996. Juvenile rheumatoid arthritis in velo-cardio-facial syndrome: coincidence or unusual complication? *Am. J. Med. Genet.* 64, 546-550.

307. Rathi,A., Virmani,A.K., Harada,K., Timmons,C.F., Miyajima,K., Hay,R.J., Mastrangelo,D., Maitra,A., Tomlinson,G.E. & Gazdar,A.F. 2003. Aberrant methylation of the HIC1 promoter is a frequent event in specific pediatric neoplasms. *Clin. Cancer Res.* 9, 3674-3678.

308. Rauch,A., Pfieffer,R.A., Leipold,G., Singer,H., Tigges,M. & Hofbeck,M. 1999. A novel 22q11.2 microdeletion in DiGeorge syndrome. *Am. J. Hum. Genet.* 64, 658-666.

309. Rauch,A., Zink,S., Zweier,C., Thiel,C.T., Koch,A., Rauch,R., Lascorz,J., Huffmeier,U., Weyand,M., Singer,H. & Hofbeck,M. 2005. Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *J. Med. Genet.*

310. Ravnan,J.B., Tepperberg,J.H., Papenhausen,P., Lamb,A.N., Hedrick,J., Eash,D., Ledbetter,D.H. & Martin,C.L. 2005. Subtelomere FISH analysis of 11,688 cases: An evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J Med. Genet.*

311. Ray-Gallet,D., Quivy,J.P., Scamps,C., Martini,E.M., Lipinski,M. & Almouzni,G. 2002. HIRA Is Critical for a Nucleosome Assembly Pathway Independent of DNA Synthesis. *Mol. Cell* 9, 1091-1100.

312. Rayala,S.K., den Hollander,P., Balasenthil,S., Yang,Z., Broaddus,R.R. & Kumar,R. 2005. Functional regulation of oestrogen receptor pathway by the dynein light chain 1. *EMBO Rep.* 6, 538-544.

313. Redon,R., Hussenet,T., Bour,G., Caulee,K., Jost,B., Muller,D., Abecassis,J. & du,M.S. 2002. Amplicon mapping and transcriptional analysis pinpoint cyclin L as a candidate oncogene in head and neck cancer. *Cancer Res.* 62, 6211-6217.

314. Reiss,J., Dorche,C., Stallmeyer,B., Mendel,R.R., Cohen,N. & Zabot,M.T. 1999. Human molybdopterin synthase gene: genomic structure and mutations in molybdenum cofactor deficiency type B. *Am. J. Hum. Genet.* 64, 706-711.

315. Rekdal,C., Sjottem,E. & Johansen,T. 2000. The nuclear factor SPBP contains different functional domains and stimulates the activity of various transcriptional activators. *J Biol. Chem.* 275, 40288-40300.

316. Ricchetti,E.T., States,L., Hosalkar,H.S., Tamai,J., Maisenbacher,M., McDonald-McGinn,D.M., Zackai,E.H. & Drummond,D.S. 2004. Radiographic study of the upper cervical spine in the 22q11.2 deletion syndrome. *J Bone Joint Surg. Am* 86-A, 1751-1760.

317. Ririe,K.M., Rasmussen,R.P. & Wittwer,C.T. 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Analytical Biochemistry* 245, 154-160.

318. Ritchie,K.J., Hahn,C.S., Kim,K.I., Yan,M., Rosario,D., Li,L., de la Torre,J.C. & Zhang,D.E. 2004. Role of ISG15 protease UBP43 (USP18) in innate immunity to viral infection. *Nat. Med.* 10, 1374-1378.

319. Roberts,C., Daw,S.C., Halford,S. & Scambler,P.J. 1997. Cloning and developmental expression analysis of chick Hira (Chira), a candidate gene for DiGeorge syndrome. *Human Molecular Genetics* 6, 237-245.

320. Roberts,C., Sutherland,H.F., Farmer,H., Kimber,W., Halford,S., Carey,A.H., Brickman,J.M., Wynshaw-Boris,A. & Scambler,P.J. 2002. Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. *Molecular and Cellular Biology* 22, 2318-2328.

321. Robinson,H.B.Jr. 1975. DiGeorge's or the III-IV pharyngeal pouch syndrome. Pathology and a theory of pathogenesis. *Perpectives in ped path* 2, 173-206.

322. Rodova,M., Islam,M.R., Maser,R.L. & Calvet,J.P. 2002. The polycystic kidney disease-1 promoter is a target of the beta-catenin/T-cell factor pathway. *Journal of Biological Chemistry* 277, 29577-29583.

323. Romeih,M., Cui,J., Michaille,J.J., Jiang,W. & Zile,M.H. 2003. Function of RARgamma and RARalpha2 at the initiation of retinoid signaling is essential for avian embryo survival and for distinct events in cardiac morphogenesis. *Developmental Dynamics* 228, 697-708.

324. Ruiter,E.M., Bongers,E.M., Smeets,D.F., Kuijpers-Jagtman,A.M. & Hamel,B.C. 2003. No justification of routine screening for 22q11 deletions in patients with overt cleft palate. *Clin. Genet* 64, 216-219.

325. Ruvinsky,I., Silver,L.M. & Ho,R.K. 1998. Characterization of the zebrafish tbx16 gene and evolution of the vertebrate T-box family. *Dev. Genes Evol.* 208, 94-99.

326. Ryan,A.K., Goodship,J.A., Wilson,D.I., Philip,N., Levy,A., Siedel,H., Schuffenhauer,S., Oechsler,H., Belohradsky,B., Priur,M., Aurias,A., Raymond,F.L., Clayton-Smith,J., Hatchwell,E., McKeown,C., Beemer,F.A., Dallapiccola,B., Novelli,G., Hurst,J., Ignatius,J., Green,A.J., Winter,R.M., Breuton,L., Brondum-Neilsen,K., Stewart,F., Van Essen,T., Patton,M., Patterson,J. & Scambler,P.J. 1997. Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *J. Med. Genet.* 34, 798-804.

327. Saitta,S.C., Harris,S.E., Gaeth,A.P., Driscoll,D.A., McDonald-McGinn,D.M., Maisenbacher,M.K., Yersak,J.M., Chakraborty,P.K., Hacker,A.M., Zackai,E.H., Ashley,T. & Emanuel,B.S. 2004. Aberrant interchromosomal exchanges are the predominant cause of the 22q11.2 deletion. *Human Molecular Genetics* 13, 417-428.

328. Sanchez-Martin,M., Rodriguez-Garcia,A., Perez-Losada,J., Sagrera,A., Read,A.P. & Sanchez-Garcia,I. 2002. SLUG (SNAI2) deletions in patients with Waardenburg disease. *Human Molecular Genetics* 11, 3231-3236.

329. Sandrin-Garcia,P., Macedo,C., Martelli,L.R., Ramos,E.S., Guion-Almeida,M.L., Richieri-Costa,A. & Passos,G.A. 2002. Recurrent 22q11.2 deletion in a sibship suggestive of parental germline mosaicism in velocardiofacial syndrome. *Clinical Genetics* 61, 380-383.

330. Sansal,I., Dupont,E., Toru,D., Evrard,C. & Rouget,P. 2000. NPDC-1, a regulator of neural cell proliferation and differentiation, interacts with E2F-1, reduces its binding to DNA and modulates its transcriptional activity. *Oncogene* 19, 5000-5009.

331. Saran,N.G., Pletcher,M.T., Natale,J.E., Cheng,Y. & Reeves,R.H. 2003. Global disruption of the cerebellar transcriptome in a Down syndrome mouse model. *Hum Mol. Genet* 12, 2013-2019.

332. Scambler,P., Roberts,C., Sutherland,H., Kimber,W., Lui,V., Halford,S., McKie,J., Snoeren,C.A.S., Lohman,F., Meijers,C. & Wynshaw-Boris,A. Hira, a gene from the DGS/VCFS region, is required for normal embryogenesis. Am.J.Hum.Genet. 63, A7. 1988.

333. Scambler,P.J., Kelly,D., Williamson,R., Goldberg,R. & Shprintzen,R. 1992. The velo-cardio-facial syndrome is associated with chromosome 22 deletions which encompass the DiGeorge syndrome locus. *Lancet* 339, 1138-1139.

334. Schedl,A., Ross,A., Lee,M., Engelkamp,D., Rashbass,P., van,H., V & Hastie,N.D. 1996. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* 86, 71-82.

335. Scheidler,S., Fredericks,W.J., Rauscher,F.J., III, Barr,F.G. & Vogt,P.K. 1996. The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. *Proc. Natl. Acad. Sci. U. S. A* 93, 9805-9809.

336. Schilham,M.W., Oosterwegel,M.A., Moerer,P., Ya,J., de Boer,P.A., van de,W.M., Verbeek,S., Lamers,W.H., Kruisbeek,A.M., Cumano,A. & Clevers,H. 1996. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 380, 711-714.

337. Schinzel,A., Schmid,W., Fraccaro,M., Tiepolo,L., Zuffardi,O., Opitz,J.M., Lindsten,J., Zetterqvist,P., Enell,H., Baccichetti,C., Tenconi,R. & Pagon,R.A. 1981. The 'cat eye syndrome': dicentric small marker chromosome probably derived from a no. 22 (tetrasomy 22pter-to-q11) associated with a characteristic phenotype: report of 11 patients and deliniation of the clinical picture. *Hum. Genet.* 57, 148-158.

338. Schleiffarth,J., Martinsen,B., Person,A., Sukovich D, Lohr J, Cornfield D, Ekker S & Petryk A . Wnt5a is required for the septation of the cardiac outflow tract in mice. Abstract, Weinstein Cardiovascular Development Conference . 2005.

339. Schmidt,C., Stoeckelhuber,M., McKinnell,I., Putz,R., Christ,B. & Patel,K. 2004. Wnt 6 regulates the epithelialisation process of the segmental plate mesoderm leading to somite formation. *Dev. Biol.* 271, 198-209.

340. Schneider,R.A., Hu,D., Rubenstein,J.L., Maden,M. & Helms,J.A. 2001. Local retinoid signaling coordinates forebrain and facial morphogenesis by maintaining FGF8 and SHH. *Development* 128, 2755-2767.

341. Schouten,J.P., McElgunn,C.J., Waaijer,R., Zwijnenburg,D., Diepvens,F. & Pals,G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic acids researchX* 30, e57.

342. Schub,T. 2002. The year of the flood: Tropical Storm Allison's impact on Texas Medical Center. *Lab Anim (NY)* 31, 34-39.

343. Schubert,F.R., Tremblay,P., Mansouri,A., Faisst,A.M., Kammandel,B., Lumsden,A., Gruss,P. & Dietrich,S. 2001. Early mesodermal phenotypes in splotch suggest a role for Pax3 in the formation of epithelial somites. *Developmental Dynamics* 222, 506-521.

344. SEDLACKOVA,E. 1955. [Insufficiency of palatolaryngeal passage as a developmental disorder.]. *Cas. Lek. Cesk.* 94, 1304-1307.

345. Sekine,K., Ohuchi,H., Fujiwara,M., Yamasaki,M., Yoshizawa,T., Sato,T., Yagishita,N., Matsui,D., Koga,Y., Itoh,N. & Kato,S. 1999. Fgf10 is essential for limb and lung formation. *Nat. Genet.* 21, 138-141.

346. Serth,K., Schuster-Gossler,K., Cordes,R. & Gossler,A. 2003. Transcriptional oscillation of lunatic fringe is essential for somitogenesis. *Genes Dev.* 17, 912-925.

347. Shaikh,T.H., Kurahashi,H. & Emanuel,B.S. 2001. Evolutionarily conserved low copy repeats (LCRs) in 22q11 mediate deletions, duplications, translocations, and genomic instability: an update and literature review. *Genet. Med.* 3, 6-13.

348. Shaw,C.J., Bi,W. & Lupski,J.R. 2002. Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am. J. Hum. Genet.* 71, 1072-1081.

349. Shaw-Smith,C., Redon,R., Rickman,L., Rio,M., Willatt,L., Fiegler,H., Firth,H., Sanlaville,D., Winter,R., Colleaux,L., Bobrow,M. & Carter,N.P. 2004. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J. Med. Genet.* 41, 241-248.

350. Shi,M.M. 2001. Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.* 47, 164-172.

351. Shi,Y.R., Hsieh,K.S., Wu,J.Y., Lee,C.C., Tsai,C.H., Yu,M.T., Chang,J.S. & Tsai,F.J. 2003. Genetic analysis of chromosome 22q11.2 markers in congenital heart disease. *J. Clin. Lab Anal.* 17, 28-35.

352. Shimomura,H., Sanke,T., Hanabusa,T., Tsunoda,K., Furuta,H. & Nanjo,K. 2000. Nonsense mutation of islet-1 gene (Q310X) found in a type 2 diabetic patient with a strong family history. *Diabetes* 49, 1597-1600.

353. Shprintzen,R.J., Goldberg,R.B., Lewin,M.L., Sidoti,E.J., Berkman,M.D., Argamaso,R.V. & Young,D. 1978. A new syndrome involving cleft palate, cardiac anomalies, typical facies, and learning disabilities: velo-cardio-facial syndrome. *Cleft Palate J.* 15, 56-62.

354. Sinibaldi,L., De Luca,A., Bellacchio,E., Conti,E., Pasini,A., Paloscia,C., Spalletta,G., Caltagirone,C., Pizzuti,A. & Dallapiccola,B. 2004. Mutations of the Nogo-66 receptor (RTN4R) gene in schizophrenia. *Hum. Mutat.* 24, 534-535.

355. Smith,J. 1999. T-box genes: what they do and how they do it. *Trends in Genetics* 15, 154-158.

356. Sock,E., Rettig,S.D., Enderich,J., Bosl,M.R., Tamm,E.R. & Wegner,M. 2004. Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Mol. Cell Biol.* 24, 6635-6644.

357. Solinas-Toldo,S., Lampel,S., Stilgenbauer,S., Nickolenko,J., Benner,A., Dohner,H., Cremer,T. & Lichter,P. 1997. Matrix-based comparative genomic

hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes. Cancer* 20, 399-407.

358. Solot,C.B., Gerdes,M., Kirschner,R.E., McDonald-McGinn,D.M., Moss,E., Woodin,M., Aleman,D., Zackai,E.H. & Wang,P.P. 2001. Communication issues in 22q11.2 deletion syndrome: children at risk. *Genet. Med.* 3, 67-71.

359. Sparrow,D.B., Chapman,G., Wouters,M.A., Whittock,N.V., Ellard,S., Fatkin,D., Turnpenny,P.D., Kusumi,K., Sillence,D. & Dunwoodie,S.L. 2006. Mutation of the LUNATIC FRINGE gene in humans causes spondylocostal dysostosis with a severe vertebral phenotype. *Am J Hum Genet* 78, 28-37.

360. Sreekantaiah,C., Kronn,D., Marinescu,R.C., Goldin,B. & Overhauser,J. 1999. Characterization of a complex chromosomal rearrangement in a patient with a typical catlike cry and no other clinical findings of cri-du-chat syndrome. *Am J Med. Genet* 86, 264-268.

361. Srivastava,D., Thomas,T., Lin,Q., Kirby,M.L., Brown,D. & Olson,E.N. 1997. Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nature Genetics* 16, 154-160.

362. Stalmans,I., Lambrechts,D., De Smet,F., Jansen,S., Wang,J., Maity,S., Kneer,P., Von Der,O.M., Swillen,A., Maes,C., Gewillig,M., Molin,D.G., Hellings,P., Boetel,T., Haardt,M., Compernolle,V., Dewerchin,M., Plaisance,S., Vlietinck,R., Emanuel,B., Gittenberger-de Groot,A.C., Scambler,P., Morrow,B., Driscol,D.A., Moons,L., Esguerra,C.V., Carmeliet,G., Behn-Krappa,A., Devriendt,K., Collen,D., Conway,S.J. & Carmeliet,P. 2003. VEGF: A modifier of the del22q11 (DiGeorge) syndrome? *Nat. Med.* 9, 173-182.

363. Stennard,F.A., Costa,M.W., Elliott,D.A., Rankin,S., Haast,S.J., Lai,D., McDonald,L.P., Niederreither,K., Dolle,P., Bruneau,B.G., Zorn,A.M. & Harvey,R.P. 2003. Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. *Dev. Biol.* 262, 206-224.

364. Stennard,F.A. & Harvey,R.P. 2005. T-box transcription factors and their roles in regulatory hierarchies in the developing heart. *Development* 132, 4897-4910.

365. Stephanou,A., Brar,B.K., Knight,R.A. & Latchman,D.S. 2000. Opposing actions of STAT-1 and STAT-3 on the Bcl-2 and Bcl-x promoters. *Cell Death. Differ.* 7, 329-330.

366. Stoffel,M., Karayiorgou,M., Espinosa,R., III & Beau,M.M. 1996. The human mitochondrial citrate transporter gene (SLC20A3) maps to chromosome band 22q11 within a region implicated in DiGeorge syndrome, velo-cardio-facial syndrome and schizophrenia. *Hum. Genet.* 98, 113-115.

367. Stoller,J.Z. & Epstein,J.A. 2005. Identification of a novel nuclear localization signal in Tbx1 that is deleted in DiGeorge syndrome patients harboring the 1223delC mutation. *Hum Mol. Genet* 14, 885-892.

368. Su,Q., Mochida,S., Tian,J.H., Mehta,R. & Sheng,Z.H. 2001. SNAP-29: a general SNARE protein that inhibits SNARE disassembly and is implicated in synaptic transmission. *Proc. Natl. Acad. Sci. U. S. A* 98, 14038-14043.

369. Sutherland,H.F., Wadey,R., McKie,J.M., Taylor,C., Atif,U., Johnstone,K.A., Halford,S., Kamath,S., Kim,U.-J., Goodship,J.A., Baldini,A. & Scambler,P.J. 1996. Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am. J. Hum. Genet.* 59, 23-31.

370. Tachibana,M., Kobayashi,Y. & Matsushima,Y. 2003. Mouse models for four types of Waardenburg syndrome. *Pigment Cell Res.* 16, 448-454.

371. Taddei,I., Morishima,M., Huynh,T. & Lindsay,E.A. 2001. Genetic factors are major determinants of phenotypic variability in a mouse model of the DiGeorge/del22q11 syndromes. *Proc Natl Acad Sci U S A* 98, 11428-11431.

372. Tago,K., Nakamura,T., Nishita,M., Hyodo,J., Nagai,S., Murata,Y., Adachi,S., Ohwada,S., Morishita,Y., Shibuya,H. & Akiyama,T. 2000. Inhibition of Wnt signaling by ICAT, a novel beta-catenin-interacting protein. *Genes Dev.* 14, 1741-1749.

373. Tajbakhsh,S., Rocancourt,D., Cossu,G. & Buckingham,M. 1997. Redefining the genetic hierachies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* 89, 127-138.

374. Tanaka,M., Chen,Z., Bartunkova,S., Yamasaki,N. & Izumo,S. 1999. The cardiac homeobox gene CSx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development* 126, 1269-1280.

375. Tanaka,M. & Tickle,C. 2004. Tbx18 and boundary formation in chick somite and wing development. *Dev. Biol.* 268, 470-480.

376. Tanaka,S.S., Toyooka,Y., Akasu,R., Katoh-Fukui,Y., Nakahara,Y., Suzuki,R., Yokoyama,M. & Noce,T. 2000. The mouse homolog of Drosophila Vasa is required for the development of male germ cells. *Genes Dev.* 14, 841-853.

377. Taricani,L., Tejada,M.L. & Young,P.G. 2002. The fission yeast ES2 homologue, Bis1, interacts with the Ish1 stress-responsive nuclear envelope protein. *Journal of Biological Chemistry* 277, 10562-10572.

378. Tassabehji,M., Newton,V.E. & Read,A.P. 1994. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nature Genetics* 8, 251-255.

379. Tautz,D. & Pfeifle,C. 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* 98, 81-85.

380. Thomas,J.A. & Graham,J.M., Jr. 1997. Chromosomes 22q11 deletion syndrome: an update and review for the primary pediatrician. *Clin. Pediatr. (Phila)* 36, 253-266.

381. Toby,G.G. & Golemis,E.A. 2001. Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. *Methods* 24, 201-217.

382. Touraine,R.L., Attie-Bitach,T., Manceau,E., Korsch,E., Sarda,P., Pingault,V., Encha-Razavi,F., Pelet,A., Auge,J., Nivelon-Chevallier,A., Holschneider,A.M.,

Munnes,M., Doerfler,W., Goossens,M., Munnich,A., Vekemans,M. & Lyonnet,S. 2000. Neurological phenotype in Waardenburg syndrome type 4 correlates with novel SOX10 truncating mutations and expression in developing brain. *Am. J. Hum. Genet.* 66, 1496-1503.

383. Toyofuku,T., Hong,Z., Kuzuya,T., Tada,M. & Hori,M. 2000. Wnt/frizzled-2 signaling induces aggregation and adhesion among cardiac myocytes by increased cadherin-beta-catenin complex. *JCB* 150, 225-241.

384. Tremblay,P., Dietrich,S., Mericskay,M., Schubert,F.R., Li,Z. & Paulin,D. 1998. A crucial role for Pax3 in the development of the hypaxial musculature and the long-range migration of muscle precursors. *Dev. Biol.* 203, 49-61.

385. Tremblay,P., Kessel,M. & Gruss,P. 1995. A transgenic neuroanatomical marker identifies cranial neural crest deficiencies associated with the Pax3 mutant Splotch. *Dev. Biol.* 171, 317-329.

386. Trieu,M., Ma,A., Eng,S.R., Fedtsova,N. & Turner,E.E. 2003. Direct autoregulation and gene dosage compensation by POU-domain transcription factor Brn3a. *Development* 130, 111-121.

387. Tsai,C.H., Van Dyke,D.L. & Feldman,G.L. 1999. Child with velocardiofacial syndrome and del (4)(q34.2): another critical region associated with a velocardiofacial syndrome-like phenotype [see comments]. *Am J Med Genet* 82, 336-339.

388. Tucker,A.S., Yamada,G., Grigoriou,M., Pachnis,V. & Sharpe,P.T. 1999. Fgf-8 determines rostral-caudal polarity in the first branchial arch. *Development* 126, 51-61.

389. van den,H.L., Ruitenbeek,W., Smeets,R., Gelman-Kohan,Z., Elpeleg,O., Loeffen,J., Trijbels,F., Mariman,E., de Bruijn,D. & Smeitink,J. 1998. Demonstration of a new pathogenic mutation in human complex I deficiency: a 5-bp duplication in the nuclear gene encoding the 18-kD (AQDQ) subunit. *Am. J. Hum. Genet.* 62, 262-268.

390. Van Esch,H., Groenen,P., Nesbit,M.A., Schuffenhauer,S., Lichtner,P., Vanderlinden,G., Harding,B., Beetz,R., Bilous,R.W., Holdaway,I., Shaw,N.J., Fryns,J.P., Van,d., V, Thakker,R.V. & Devriendt,K. 2000. GATA3 haplo-insufficiency causes human HDR syndrome. *Nature* 406, 419-422.

391. Van Geet,C., Devriendt,K., Eyskens,B., Vermylen,J. & Hoylaerts,M.F. 1998. Velocardiofacial syndrome patients with a heterozygous chromosome 22q11 deletion have giant platelets. *Pediatr. Res.* 44, 607-611.

392. Vandesompele,J., De Paepe,A. & Speleman,F. 2002. Elimination of primer-dimer artifacts and genomic coamplification using a two-step SYBR green I real-time RT-PCR. *Analytical Biochemistry* 303, 95-98.

393. Velculescu,V.E., Zhang,L., Vogelstein,B. & Kinzler,K.W. 1995. Serial analysis of gene expression. *Science* 270, 484-487.

394. Veltman,J.A., Fridlyand,J., Pejavar,S., Olshen,A.B., Korkola,J.E., DeVries,S., Carroll,P., Kuo,W.L., Pinkel,D., Albertson,D., Cordon-Cardo,C., Jain,A.N. &

Waldman,F.M. 2003. Array-based comparative genomic hybridization for genome-wide screening of DNA copy number in bladder tumors. *Cancer Res.* 63, 2872-2880.

395. Veltman,J.A., Schoenmakers,E.F., Eussen,B.H., Janssen,I., Merkx,G., van Cleef,B., van Ravenswaaij,C.M., Brunner,H.G., Smeets,D. & Van Kessel,A.G. 2002. High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. *Am. J. Hum. Genet.* 70, 1269-1276.

396. Vermot,J., Niederreither,K., Garnier,J.M., Chambon,P. & Dolle,P. 2003. Decreased embryonic retinoic acid synthesis results in a DiGeorge syndrome phenotype in newborn mice. *Proc. Natl. Acad. Sci. U. S. A* 100, 1763-1768.

397. Vissers,L.E., De Vries,B.B., Osoegawa,K., Janssen,I.M., Feuth,T., Choy,C.O., Straatman,H., Van,D., V, Huys,E.H., Van Rijk,A., Smeets,D., Ravenswaaij-Arts,C.M., Knoers,N.V., van,d.B., I, De Jong,P.J., Brunner,H.G., Van Kessel,A.G., Schoenmakers,E.F. & Veltman,J.A. 2003. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am. J. Hum. Genet.* 73, 1261-1270.

398. Vissers,L.E., van Ravenswaaij,C.M., Admiraal,R., Hurst,J.A., De Vries,B.B., Janssen,I.M., van der Vliet,W.A., Huys,E.H., De Jong,P.J., Hamel,B.C., Schoenmakers,E.F., Brunner,H.G., Veltman,J.A. & Van Kessel,A.G. 2004. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet* 36, 955-957.

399. Vissers,L.E., Veltman,J.A., Van Kessel,A.G. & Brunner,H.G. 2005. Identification of disease genes by whole genome CGH arrays. *Hum Mol. Genet* 14 Spec No. 2, R215-R223.

400. Vitelli,F., Morishima,M., Taddei,I., Lindsay,E.A. & Baldini,A. 2002a. Tbx1 mutation causes multiple cardiovascular defects and disrupts neural crest and cranial nerve migratory pathways. *Human Molecular Genetics* 11, 915-922.

401. Vitelli,F., Taddei,I., Morishima,M., Meyers,E.N., Lindsay,E.A. & Baldini,A. 2002b. A genetic link between Tbx1 and fibroblast growth factor signaling. *Development* 129, 4605-4611.

402. Volpe,P., Marasini,M., Caruso,G., Marzullo,A., Buonadonna,A.L., Arciprete,P., Di Paolo,S., Volpe,G. & Gentile,M. 2003. 22q11 deletions in fetuses with malformations of the outflow tracts or interruption of the aortic arch: impact of additional ultrasound signs. *Prenat. Diagn.* 23, 752-757.

403. Waibler,Z., Schafer,A. & Starzinski-Powitz,A. 2001. mARVCF cellular localisation and binding to cadherins is influenced by the cellular context but not by alternative splicing. *J. Cell Sci.* 114, 3873-3884.

404. Wang,W., Lo,P., Frasch,M. & Lufkin,T. 2000. Hmx: an evolutionary conserved homeobox gene family expressed in the developing nervous system in mice and Drosophila. *Mechanisms of Development* 99, 123-137.

405. Watanabe,A., Takeda,K., Ploplis,B. & Tachibana,M. 1998. Epistatic relationship between Waardenburg syndrome genes *MITF* and *PAX3*. *Nature Genetics* 18, 283-286.

406. Weinzimer,S.A. & Driscoll,D.A. 2001. Endocrine aspects of the 22q11.2 deletion syndrome

407. Wendling,O., Dennefeld,C., Chambon,P. & Mark,M. 2000. Retinoid signaling is essential for patterning the endoderm of the third and fourth pharyngeal arches. *Development* 127, 1553-1562.

408. Wendling,O., Ghyselinck,N.B., Chambon,P. & Mark,M. 2001. Roles of retinoic acid receptors in early embryonic morphogenesis and hindbrain patterning. *Development* 128, 2031-2038.

409. Wessendorf,S., Schwaenen,C., Kohlhammer,H., Kienle,D., Wrobel,G., Barth,T.F., Nessling,M., Moller,P., Dohner,H., Lichter,P. & Bentz,M. 2003. Hidden gene amplifications in aggressive B-cell non-Hodgkin lymphomas detected by microarray-based comparative genomic hybridization. *Oncogene* 22, 1425-1429.

410. White,P.H., Farkas,D.R., McFadden,E.E. & Chapman,D.L. 2003. Defective somite patterning in mouse embryos with reduced levels of Tbx6. *Development* 130, 1681-1690.

411. Wilhelm,M., Veltman,J.A., Olshen,A.B., Jain,A.N., Moore,D.H., Presti,J.C., Jr., Kovacs,G. & Waldman,F.M. 2002. Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer. *Cancer Res.* 62, 957-960.

412. Wilkinson,D.G., Bhatt,S. & Herrmann,B.G. 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* 343, 657-659.

413. Williams,B.A. & Ordahl,C.P. 1994. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. *Development* 120, 785-796.

414. Wilming,L.G., Snoeren,C.A.S., van Rijswijk,A., Grosveld,F. & Meijers,C. 1997. The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in CATCH22 patients. *Human Molecular Genetics* 6, 247-258.

415. Wilson,D.I., Cross,I.E., Wren,C., Scambler,P.J., Burn,J. & Goodship,J. Minimum prevalence of chromosome 22q11 deletions. Am.J.Hum.Genet. 55, A169. 1994.

416. Wilson,M.J., Salata,M.W., Susalka,S.J. & Pfister,K.K. 2001. Light chains of mammalian cytoplasmic dynein: identification and characterization of a family of LC8 light chains. *Cell Motil. Cytoskeleton* 49, 229-240.

417. Winer,J., Jung,C.K., Shackel,I. & Williams,P.M. 1999. Development and validation of real-time quantitative reverse transcriptase-polymerase chain

reaction for monitoring gene expression in cardiac myocytes in vitro. *Analytical Biochemistry* 270, 41-49.

418. Winnier,G., Blessing,M., Labosky,P.A. & Hogan,B.L. 1995. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev.* 9, 2105-2116.

419. Woodin,M., Wang,P.P., Aleman,D., McDonald-McGinn,D., Zackai,E. & Moss,E. 2001. Neuropsychological profile of children and adolescents with the 22q11.2 microdeletion. *Genet. Med.* 3, 34-39.

420. Woods,C.G., Valente,E.M., Bond,J. & Roberts,E. 2004. A new method for autozygosity mapping using single nucleotide polymorphisms (SNPs) and EXCLUDEAR. *J. Med. Genet.* 41, e101.

421. Wu,H.Y., Rusnack,S.L., Bellah,R.D., Plachter,N., McDonald-McGinn,D.M., Zackai,E.H. & Canning,D.A. 2002. Genitourinary malformations in chromosome 22q11.2 deletion. *J. Urol.* 168, 2564-2565.

422. Wu,M., Hemesath,T.J., Takemoto,C.M., Horstmann,M.A., Wells,A.G., Price,E.R., Fisher,D.Z. & Fisher,D.E. 2000. c-Kit triggers dual phosphorylations, which couple activation and degradation of the essential melanocyte factor Mi. *Genes Dev.* 14, 301-312.

423. Xu,H., Morishima,M., Wylie,J.N., Schwartz,R.J., Bruneau,B.G., Lindsay,E.A. & Baldini,A. 2004. Tbx1 has a dual role in the morphogenesis of the cardiac outflow tract. *Development* 131, 3217-3227.

424. Yagi,H., Furutani,Y., Hamada,H., Sasaki,T., Asakawa,S., Minoshima,S., Ichida,F., Joo,K., Kimura,M., Imamura,S., Kamatani,N., Momma,K., Takao,A., Nakazawa,M., Shimizu,N. & Matsuoka,R. 2003. Role of TBX1 in human del22q11.2 syndrome. *Lancet* 362, 1366-1373.

425. Yamagishi,C., Hierck,B.P., Gittenberger-de Groot,A.C., Yamagishi,H. & Srivastava,D. 2003a. Functional attenuation of UFD1l, a 22q11.2 deletion syndrome candidate gene, leads to cardiac outflow septation defects in chicken embryos. *Pediatr. Res.* 53, 546-553.

426. Yamagishi,H., Garg,V., Matsuoka,R., Thomas,T. & Srivastava,D. 1999. A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. *Science* 283, 1158-1161.

427. Yamagishi,H., Maeda,J., Hu,T., McAnally,J., Conway,S.J., Kume,T., Meyers,E.N., Yamagishi,C. & Srivastava,D. 2003b. Tbx1 is regulated by tissue-specific forkhead proteins through a common Sonic hedgehog-responsive enhancer. *Genes Dev.* 17, 269-281.

428. Yan,W., Jacobsen,L.K., Krasnewich,D.M., Guan,X.Y., Lenane,M.C., Paul,S.P., Dalwadi,H.N., Zhang,H., Long,R.T., Kumra,S., Martin,B.M., Scambler,P.J., Trent,J.M., Sidransky,E., Ginns,E.I. & Rapoport,J.L. 1998. Chromosome 22q11.2 interstitial deletions among childhood-onset schizophrenics and "multidimensionally impaired". *Am J Med Genet* 81, 41-43.

429. Ye,G., Donthi,R.V., Metreveli,N.S. & Epstein,P.N. 2005. Cardiomyocyte dysfunction in models of type 1 and type 2 diabetes. *Cardiovasc. Toxicol.* 5, 285-292.

430. Yi,C.H., Terrett,J.A., Li,Q.Y., Ellington,K., Packham,E.A., Armstrong-Buisseret,L., McClure,P., Slingsby,T. & Brook,J.D. 1999. Identification, mapping, and phylogenomic analysis of four new human members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19. *Genomics* 55, 10-20.

431. Yobb,T.M., Somerville,M.J., Willatt,L., Firth,H.V., Harrison,K., MacKenzie,J., Gallo,N., Morrow,B.E., Shaffer,L.G., Babcock,M., Chernos,J., Bernier,F., Sprysak,K., Christiansen,J., Haase,S., Elyas,B., Lilley,M., Bamforth,S. & McDermid,H.E. 2005. Microduplication and triplication of 22q11.2: a highly variable syndrome. *Am. J. Hum. Genet.* 76, 865-876.

432. Yoshida,K., Kuo,F., George,E.L., Sharpe,A.H. & Dutta,A. 2001. Requirement of CDC45 for postimplantation mouse development. *Mol. Cell Biol.* 21, 4598-4603.

433. Yoshida,K., Taga,T., Saito,M., Suematsu,S., Kumanogoh,A., Tanaka,T., Fujiwara,H., Hirata,M., Yamagami,T., Nakahata,T., Hirabayashi,T., Yoneda,Y., Tanaka,K., Wang,W.Z., Mori,C., Shiota,K., Yoshida,N. & Kishimoto,T. 1996. Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc. Natl. Acad. Sci. U. S. A* 93, 407-411.

434. Yu,J., Yu,L., Chen,Z., Zheng,L., Chen,X., Wang,X., Ren,D. & Zhao,S. 2002. Protein inhibitor of neuronal nitric oxide synthase interacts with protein kinase A inhibitors. *Brain Res. Mol. Brain Res.* 99, 145-149.

435. Yu,W., Ballif,B.C., Kashork,C.D., Heilstedt,H.A., Howard,L.A., Cai,W.W., White,L.D., Liu,W., Beaudet,A.L., Bejjani,B.A., Shaw,C.A. & Shaffer,L.G. 2003. Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Human Molecular Genetics* 12, 2145-2152.

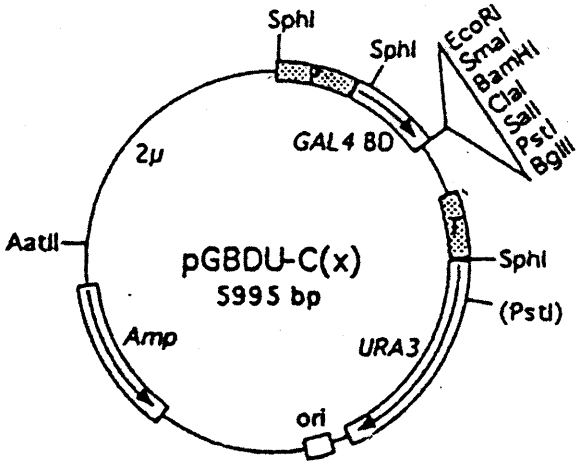
436. Zackai,E.H. & Emanuel,B.S. 1980. Site-specific reciprocal translocation, t(11;22) (q23;q11), in several unrelated families with 3:1 meiotic disjunction. *Am. J. Med. Genet.* 7, 507-521.

437. Zemmyo,M., Meharra,E.J., Kuhn,K., Creighton-Achermann,L. & Lotz,M. 2003. Accelerated, aging-dependent development of osteoarthritis in alpha1 integrin-deficient mice. *Arthritis Rheum.* 48, 2873-2880.

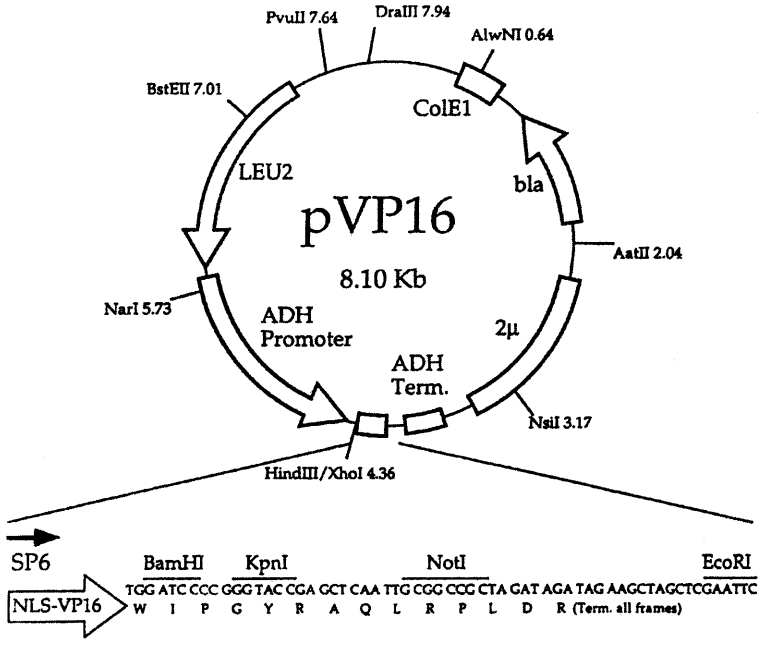
438. Zhang,R., Poustovoitov,M.V., Ye,X., Santos,H.A., Chen,W., Daganzo,S.M., Erzberger,J.P., Serebriiskii,I.G., Canutescu,A.A., Dunbrack,R.L., Pehrson,J.R., Berger,J.M., Kaufman,P.D. & Adams,P.D. 2005a. Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA. *Dev. Cell* 8, 19-30.

439. Zhang,X., Snijders,A., Segraves,R., Zhang,X., Niebuhr,A., Albertson,D., Yang,H., Gray,J., Niebuhr,E., Bolund,L. & Pinkel,D. 2005b. High-resolution mapping of genotype-phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. *Am J Hum Genet* 76, 312-326.

440. Zhao,C., Takita,J., Tanaka,Y., Setou,M., Nakagawa,T., Takeda,S., Yang,H.W., Terada,S., Nakata,T., Takei,Y., Saito,M., Tsuji,S., Hayashi,Y. & Hirokawa,N. 2001. Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. *Cell* 105, 587-597.



PGAD-C3 AAA GAG ATC GAA TTC CCG GGG* GAT CCA TCG ATG TCG ACC TGC AGA GAT CTA TGA ATC GTA GAT ACT GAA



Appendix 1.

Maps of vectors used in the yeast two hybrid experiment. Polylinker translational reading frame: restriction sites are underlined, stop codons are boxed, * single base pair insertions introduced during construction.

Affy No	Gene	P value
160552	adaptor protein complex AP-1, sigma 1	0.0000102
97966	hypothetical protein MGC28864	0.0000102
96174	nuclear pore membrane protein 121	5.544E-05
101898	unnamed protein product; Mouse MHC (Qa) Q10-k gene for class I antigen.	5.544E-05
100965	RIKEN cDNA 1700007F21 gene	0.0003419
96242	similar to RIKEN cDNA 2210412D01	0.0004705
97660	Mus musculus mRNA for TCF-4 protein.	0.000502
92302	Son of sevenless homolog 2 (Drosophila)	0.0009669
96898	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1	0.001143
100440	alternatively spliced form of ankyrin; major form of ankyrin found in red blood cells	0.001566
100344	Mus musculus, similar to quinone reductase-like protein, clone IMAGE:4972406	0.001993
101685	uc45a07.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone IMAGE	0.002309
96723	synovial sarcoma, X breakpoint 2 interacting protein	0.002309
99701	Mus Musculus SPRR2B gene.	0.002309
96643	RIKEN cDNA 1600023A02 gene	0.002415
162018	crystallin, gamma F	0.002671
95184	vc51g02.r1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:778130	0.002671
101344	cholecystokinin B receptor	0.002671
160372	Mus musculus gene for LKB1 serine/threonin kinase, exon 2-10, partial cds.	0.002671
96565	ESTs	0.002671
100327	precursor; Mouse LT lymphotoxin (LT) gene, complete cds.	0.002804
98765	immunoglobulin heavy chain (J558 family)	0.002804
92979	ets variant gene 4 (E1A enhancer binding protein, E1AF)	0.002804
93693	RIKEN cDNA F830020C16 gene	0.002804
161535	polycystic kidney disease 1 homolog	0.002804
96839	RIKEN cDNA 3110057M17 gene	0.002804
161812	RIKEN cDNA 4632419I22 gene	0.002804
95104	syndecan 2	0.002804
101734	glutamate receptor, ionotropic, delta 2	0.002919
161264	RIKEN cDNA 3200002M19 gene	0.002919
103514	tumor necrosis factor receptor superfamily, member 21	0.003193
98029	RIKEN cDNA 3110056O03 gene	0.003477
93187	RIKEN cDNA 2210023F24 gene	0.003688
98808	neurogenic differentiation 2	0.003705
161641	cDNA sequence BC008155	0.003705
161201	solute carrier family 9 (sodium/hydrogen exchanger), member 1	0.004029
103552	rabaptin 5	0.004029
161859	synuclein, gamma	0.004029
103399	sex comb on midleg homolog 1	0.00403
97983	syntaxin binding protein 1	0.00403
104255	ESTs, Weakly similar to DIA3_MOUSE Diaphanous protein homolog 3	0.004892
101650	protocadherin alpha 6	0.005893
97224	UI-M-BH0-akh-h-10-0-UI.s1 NIH_BMAP_M1 Mus musculus cDNA clone	0.006245
162313	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3	0.006245
93266	tropomyosin 3, gamma	0.006245
100282	5-hydroxytryptamine (serotonin) receptor 7	0.006445
161644	guanidinoacetate methyltransferase	0.006473
104170	mitogen activated protein kinase 8 interacting protein	0.0066
160629	regulator of G-protein signalling 10	0.0066
93996	cytochrome P450, family 2, subfamily e, polypeptide 1	0.00664
103205	T-cell, immune regulator 1	0.006643
162356	RIKEN cDNA 1810042K04 gene	0.006829
92437	RIKEN cDNA 3110041O18 gene	0.007075
162177	ribosomal protein L3	0.007075
92914	homeo box 2.3; putative; Mouse homeo box (Hox2.3) region.	0.007168
95948	ESTs	0.00723
93308	pyruvate carboxylase	0.007315
101856	RIKEN cDNA 2900008M13 gene	0.007315

98277	zinc finger protein, subfamily 1A, 4	0.007315
100467	lymphoblastomic leukemia	0.007324
160409	phosphatidylinositol transfer protein	0.007361
161911	RIKEN cDNA 2700069E09 gene	0.007361
100318	RIKEN cDNA 1700014P03 gene	0.007598
160895	Fanconi anemia, complementation group C	0.007598
101684	M.musculus T2 mRNA.	0.007874
101806	Mus musculus galanin receptor type 3 (GalR3) gene, complete cds.	0.008416
100717	cDNA sequence U90926	0.008416
103651	general transcription factor IIF, polypeptide 2	0.008609
104263	DNA segment, Chr 4, Wayne State University 24, expressed	0.008609
162250	ESTs	0.008642
97375	polycystic kidney disease 1 homolog	0.008642
161120	AV132572 Mus musculus C57BL/6J 11-day embryo Mus musculus cDNA clone 2700091C09	0.008654
93682	Mus musculus LIM homeobox protein cofactor CLIM-1b mRNA, complete cds.	0.008654
93073	nuclear factor of activated T-cells, cytoplasmic 2	0.008821
101546	Intron-exon boundaries defined in relation to human cDNA in L03411.; complement factor B	0.009382
99992	hypothetical protein LOC231291	0.009382
102383	RIKEN cDNA 5730593F17 gene	0.01018
102261	procollagen, type XIII, alpha 1	0.01018
160874	RuvB-like protein 1	0.01025
93648	protein kinase C, gamma	0.01034
102895	RIKEN cDNA 4921518A06 gene	0.01034
93018	interferon induced transmembrane protein 3-like	0.01062
161990	diaphanous homolog 1 (Drosophila)	0.01076
161650	secretory leukocyte protease inhibitor	0.01082
160085	thiosulfate sulfurtransferase, mitochondrial	0.01099
161446	protease, serine, 25	0.01124
102426	calsequestrin 1	0.01198
160920	Bcl2-like 2	0.01221
92635	tubulin, alpha 4	0.01221
97053	single cysteine motif 1 receptor; Mus musculus gene for SCM1 receptor mXCR1	0.01223
93409	ESTs	0.01239
100617	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5	0.01239
162014	RIKEN cDNA 1300011C24 gene	0.01239
98278	sprouty homolog 4 (Drosophila)	0.01248
94196	inhibitor of kappaB kinase gamma	0.01248
98131	crystallin, zeta	0.01252
96099	casein kinase II, beta subunit	0.0126
95883	RIKEN cDNA D530048A03 gene	0.01295
94917	F-box only protein 8	0.01338
98798	chondroitin sulfate proteoglycan 3	0.01338
92702	astrotactin 1	0.01355
99421	M.musculus Ig Vheavy-PCG-4 gene.	0.01355
161396	RIKEN cDNA B230113M03 gene	0.01431
98486	stathmin-like 4	0.01431
161626	plasminogen	0.01471
101217	RIKEN cDNA D930050G01 gene	0.01501
98926	vesicle-associated membrane protein 2	0.01525
101131	cholinergic receptor, nicotinic, alpha polypeptide 7	0.01596
103729	laminin, alpha 1	0.01752
104568	myeloid/lymphoid or mixed-lineage leukemia	0.0185
101395	purine rich element binding protein B	0.01863
161030	ESTs	0.01918
99883	iduronate 2-sulfatase	0.01918
160336	isoprenylcysteine carboxyl methyltransferase	0.01918
160134	RIKEN cDNA 2810031L11 gene	0.01918
99492	catenin beta interacting protein 1	0.01918
98031	Bcl-2-related ovarian killer protein	0.02007
101307	M.musculus cyp4a12 gene, exons 10 & 11 (joined CDS).	0.02007

101623	solute carrier family 16 (monocarboxylic acid transporters), member 8	0.02007
97398	RIKEN cDNA 9130022B02 gene	0.02033
94010	RIKEN cDNA 2610301I15 gene	0.02037
96972	put. Ig kappa precursor; Mouse DNA for Ig-kappa light chain V-J kappa 5 joining region.	0.0205
93900	UI-M-BH1-ans-d-03-0-UI.s1 NIH_BMAP_M2 Mus musculus cDNA clone	0.0205
102919	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein K (avian)	0.0205
160093	RIKEN cDNA D730042P09 gene	0.02053
99641	DNA segment, Chr 2, Brigham & Women's Genetics 0891 expressed	0.02053
97151	orphan G protein-coupled receptor	0.02098
96767	membrane bound C2 domain containing protein	0.02197
161348	PDZ and LIM domain 1 (elfin)	0.02197
102117	RIKEN cDNA 2600013G09 gene	0.02279
102963	E2F transcription factor 1	0.02338
95992	ESTs, Weakly similar to 1614337A formin [Mus musculus] [M.musculus]	0.0237
98800	presumed nucleobase permease; Description: yolk sac permease-like molecule 1 form 1	0.02422
95952	RIKEN cDNA 1810073P09 gene	0.0246
94288	Mouse histone H1 gene, complete cds.	0.0246
95466	coactosin-like 1 (Dictyostelium)	0.02488
96310	myelin basic protein	0.02503
102372	immunoglobulin joining chain	0.02692
102411	zinc finger protein 239	0.02776
99100	signal transducer and activator of transcription 3	0.02915
161545	myosin, heavy polypeptide 4, skeletal muscle	0.02915
99399	Mus musculus agouti-related protein (Agrp) gene, complete cds.	0.02946
160869	sirtuin 3 (silent mating type information regulation 2, homolog) 3 (S. cerevisiae)	0.02989
161476	Mus musculus, clone MGC:31065 IMAGE:4035973, mRNA, complete cds	0.02989
161548	receptor-like tyrosine kinase	0.02989
94709	glial cells missing homolog 2 (Drosophila)	0.02989
96737	similar to hypothetical protein FLJ20337	0.02989
103712	UI-M-BH2.1-aqa-c-12-0-UI.s1 NIH_BMAP_M3.1	0.03003
161786	integrin beta 5	0.03075
101341	Mus musculus MHC class Ib antigen (M9) gene, complete cds.	0.03191
161826	glutamate-ammonia ligase (glutamine synthase)	0.03212
104680	receptor (calcitonin) activity modifying protein 1	0.03212
99007	flotillin 2	0.03327
161573	RIKEN cDNA E430014N10 gene	0.03327
98008	chemokine (C-X3-C motif) ligand 1	0.03327
104273	bile acid-Coenzyme A: amino acid N-acyltransferase	0.03327
101049	ribosomal protein L12	0.03349
97317	ectonucleotide pyrophosphatase/phosphodiesterase 2	0.03349
162269	RIKEN cDNA 1110020E07 gene	0.03532
96968	limitin	0.03589
161500	Bcl-associated death promoter	0.03589
92235	retinoid X receptor alpha	0.03883
99847	sialyltransferase 4A (beta-galactosidase alpha-2,3-sialytransferase)	0.03919
161920	enhancer trap locus 1	0.03973
161099	hypothetical protein 9430022A14	0.04122
94385	RIKEN cDNA D530030D03 gene	0.04393
92764	kinase suppressor of ras	0.04688
93525	ubiquitin-conjugating enzyme E2, J2 homolog (yeast)	0.0483
162317	AV064697 Mus musculus small intestine C57BL/6J adult Mus musculus cDNA clone 2010011P14	0.0483
98126	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	0.0483
96184	RIKEN cDNA 1110014H17 gene	0.0483
93231	Mus musculus hic-1 gene, partial.	0.0483
102131	Mus musculus, clone IMAGE:3483400, mRNA	0.0483

Institute of Child Health

and Great Ormond Street Hospital for Children NHS Trust
UNIVERSITY COLLEGE LONDON

26 April 2001

Professor PJ Scambler
Molecular Medicine Unit
ICH

Dear Prof Scambler,

01MM01 Analysis of the DiGeorge Haploinsufficiency Syndrome

Notification of ethical approval

The above research has been given ethical approval after review by the Chairman of the Great Ormond Street Hospital for Children NHS Trust / Institute of Child Health Research Ethics Committee subject to the following conditions.

1. Your research must commence within twelve months of the date of this letter and ethical approval is given for a period of 36 months from the commencement of the project. If you wish to start the project more than twelve months from the date of this letter you should seek Chairman's approval.
2. You must seek Chairman's approval for proposed amendments to the research for which this approval has been given. Ethical approval is specific to this project and must not be treated as applicable to research of a similar nature, e.g. using the same procedure(s) or medicinal product(s). Each research project is reviewed separately and if there are significant changes to the research protocol, for example in response to a grant giving body's requirements, you should seek confirmation of continued ethical approval.
3. On completion of the research, you must submit a report of your findings to the Research Ethics Committee.
4. Researchers are reminded that REC approval does not imply approval by the GOS Trust. Researchers should confirm with the R&D office that all necessary permissions have been obtained before proceeding.

THE QUEEN'S ANNIVERSARY PRIZES
FOR HIGHER AND FURTHER EDUCATION
2000

UCL

Appendix 3.1
Research Ethical Committee approval document.

5. It is your responsibility to notify the Committee immediately of any information which would raise questions about the safety and continued conduct of the research.
6. Specific conditions pertaining to the approval of this project are:

- The use of the enclosed standard consent forms for the research. A copy of the signed consent form must be placed in the patient's clinical records and a copy must be kept by you with the research records.

Yours sincerely

Orlagh Sheils
Administrator to the Research Ethics Committee

Appendix 3.2.
Research Ethical Committee approval document.

REC No. 01MM01

Great Ormond Street Hospital for Children NHS Trust and Institute of Child Health Research Ethics Committee

Consent Form for PARENTS OR GUARDIANS of Children Participating in Research Studies

Title: Analyysis of the TB1 gene and the Df1/DiGeorge syndrome haploinsufficiency Syndromes.

NOTES FOR PARENTS OR GUARDIANS

1. Your child has been asked to take part in a research study. The person organising that study is responsible for explaining the project to you before you give consent.

2. Please ask the researcher any questions you may have about this project, before you decide whether you wish to participate.

3. If you decide, now or at any other stage, that you do not wish your child to participate in the research project, that is entirely your right, and if your child is a patient it will not in any way prejudice any present or future treatment.

4. You will be given an information sheet which describes the research project. This information sheet is for you to keep and refer to. ***Please read it carefully.***

5. If you have any complaints about the way in which this research project has been or is being conducted, please, in the first instance, discuss them with the researcher. If the problems are not resolved, or you wish to comment in any other way, please contact the Chairman of the Research Ethics Committee, by post via The Research and Development Office, Institute of Child Health, 30 Guilford Street, London WC1N 1EH or if urgent, by telephone on 020 7905 2620 and the committee administration will put you in contact with him.

CONSENT

I/We _____, being the parent(s)/guardian(s) of _____ agree that the Research Project named above has been explained to me to my/our satisfaction, and I/We give permission for our child to take part in this study. I/We have read both the notes written above and the Information Sheet provided, and understand what the research study involves.

SIGNED (Parent (s)/Guardian (s)) **DATE**

_____ _____

SIGNED (Researcher) **DATE**

_____ _____

Appendix 4.1
Patient consent form.

INFORMATION SHEET (Patient)

WHAT IS THE CAUSE OF DiGEORGE AND VELOCARDIOFACIAL (SHPRINTZEN) SYNDROMES WHEN NO CHROMOSOME 22 DELETION CAN BE FOUND?

Your doctor will have discussed with you the possibility that you have the DiGeorge or velocardiofacial (Shprintzen) syndrome. Often, these syndromes are caused by the absence of a part of one of the pair of chromosomes given the number 22; this is called a chromosome deletion, or microdeletion. The chromosomes contain the genetic information (the "genes") needed for proper development and maintenance of the human body. The deletion removes certain genes from the chromosome and in this way disturbs the process of development. At the moment no-one knows what causes these syndromes when the chromosomes appear to be intact, in other words when no genes are missing. However, scientific studies into how babies develop has provided certain hints as to what might be happening in these cases. As a consequence you will be offered an extension of the deletion (FISH) test if it turns out your chromosomes 22 are normal. These tests will research the possibility that individual genes are not working properly by checking the genes' chemical composition in yourself and possibly your parents. If any abnormality is found then your doctor will be informed and will be able advise you in the same way as he/she would if the deletion test proved positive. These tests are at the research stage and it is quite possible nothing will be found. Your participation entirely voluntary.

Appendix 4.2.
Patient information sheet

Institute of Child Health

and Great Ormond Street Hospital for Children NHS Trust

UNIVERSITY COLLEGE LONDON

Molecular Medicine Unit
Institute of Child Health
30 Guilford Street
London
WC1N 1EH
United Kingdom

03/09/03

Dear Parents/Guardians of .....

I am a research doctor looking into the causes of congenital heart defects. I have been sent DNA from ..... by....., consultant geneticist. I am writing to ask your permission to put .....'s DNA sample on a new diagnostic technology we are developing.

It is something called a microarray chip and it is very sensitive at detecting small changes in chromosome number. Chromosomes are present in every cell of our bodies and contain the genetic information necessary for us to develop into humans. They determine our eye and hair colour for instance. Sometimes small changes in chromosomal number can cause individuals to develop heart defects and other problems. It may possibly give us an answer as to why..... has the heart and other problems that he has.

If you would like to participate in the study, please return the consent forms in the enclosed stamped addressed envelope. If you agree I will send you the results and can arrange to discuss them if they are positive. If you have any queries about this letter please don't hesitate to contact me at the address/ numbers above. Many thanks for your help with this research.

Yours Sincerely,

Katrina
Dr Katrina Prescott MA MBBS MRCPCH
Clinical Research Fellow-Genetics.

Appendix 5.1
Letter to the patient's parents.

Institute of Child Health

and Great Ormond Street Hospital for Children NHS Trust

UNIVERSITY COLLEGE LONDON

Molecular Medicine Unit
Institute of Child Health
30 Guilford Street
London
WC1N 1EH
United Kingdom

CONSENT TO USE DNA ON GENOMIC MICROARRAY

Ref No:

I give permission for DNA of:

Name in Full: _____dob: __

to be used on the microarray chip. I would like my results to be given to me in due course.

Signed: _____date:

Name in Capitals: _____

Relationship to Patient:_____

Comments: _____

Appendix 5.2
Patient genomic array consent form.

Institute of Child Health

and Great Ormond Street Hospital for Children NHS Trust

UNIVERSITY COLLEGE LONDON

Molecular Medicine Unit
Institute of Child Health
30 Guilford Street
London
WC1N 1EH
United Kingdom

20/02/04

Dear

Thank you very much for sending DNA on ..... and family. The sample has been analysed on the genomic Sanger array-CGH facility and no chromosomal deletions or duplications were detected at the 1Mb resolution. We will obviously still consider the sample should new technologies become available. The sample may be screened for specific gene mutations in the future aswell.

Yours Sincerely,

Katrina Prescott MA MBBS MRCPCH
Clinical Research Fellow
Prof. Pete Scambler's Lab.

Appendix 6
Clinician results letter

Publications resulting from thesis